



SYNTHESIS OF DISTRIBUTED SYSTEMS

N69-19554
NASA CR-66745-

**CASE FILE
COPY**

**THE UNIVERSITY OF ROCHESTER
COLLEGE OF ENGINEERING AND APPLIED SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING
ROCHESTER, NEW YORK**

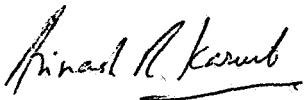
ANNUAL REPORT

NASA RESEARCH GRANT NGR-33-019-058

SYNTHESIS OF DISTRIBUTED SYSTEMS

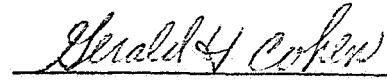
September 1, 1967 - August 31, 1968

Research Assistant



Avinash Karnik

Principal Investigator



Gerald H. Cohen

Department of Electrical Engineering
College of Engineering and Applied Science
The University of Rochester
Rochester, New York

INDEX

	page
Introduction	1
Mathematical Formulation of the Oscillator Problem	4
Methods of Solution	9
Algorithms and Problems	25
Computer Solutions	37
Errors and Limitations	53
Conclusions	55
Appendix A	56
Appendix B	64

SYMBOLS

$r(x)$	Distributed series resistance
$c(x)$	Distributed shunt capacitance
$\mathcal{L}(x)$	Distributed series inductance
$v(x, t)$	Distributed line voltage
$i(x, t)$	Distributed line current
L	Length of the line
Ω	Equation of constraint
ϕ	Criterion functional
J	Optimal criterion functional
$V(x, \omega), V(x)$	Line phasor voltage
$I(x, \omega), I(x)$	Line phasor current
y	$y^t = [y_1, y_2, y_3, y_4] = [V_1, V_2, I_1, I_2]$
λ	$\lambda^t = [\lambda_1, \lambda_2, \lambda_3, \lambda_4]$
u	$u^t = [u_1, u_2] = [r(x), c(x)]$
λ^ϕ	Adjoint system of variables corresponding to ϕ correction
λ^Ω	Adjoint system of variables corresponding to Ω correction
δu^ϕ	ϕ component of the variation in control variable u
δu^Ω	Ω component of the variation in control variable u
Φ	Fundamental matrix of system equations
Ψ	Fundamental matrix of adjoint equations
f	$f = \frac{d}{dx} y$
$\langle \lambda, f \rangle$	Inner product of λ and f vectors
H	Hamiltonian, $H = \langle \lambda, f \rangle$
H^ϕ	Hamiltonian corresponding to λ^ϕ system

H^Ω	Hamiltonian corresponding to λ^Ω system
H_u	Partial differential operation $H_u = \left[\frac{\partial H}{\partial u_1}, \frac{\partial H}{\partial u_2} \right]$
Sgn	Sign
L.H.S.	Left hand side
R.H.S.	Right hand side
Superscript t denotes transpose operation.	
AC	Analog computer
DC	Digital Computer
ADC	Analog to Digital converter
DAC	Digital to Analog converter

FIGURE CAPTIONS

		page
Fig. 1	Block Diagram for Phase Shift Oscillator	1
Fig. 2	Feedback Circuit for Phase Shift Oscillator	2
Fig. 3	Distributed Parameter Network	2
Fig. 4	Basic Elements for Transmission Line	4
Fig. 5	Solutions for Transmission Line	6
Fig. 5a	Solutions for 180° Phase Shift Line	7
Fig. 6	Trajectories for Optimal & Non-optimal Control	10
Fig. 7	Analog Set-up	28
Fig. 8	Computing Algorithm for Improved Hybrid Computing Technique	29
Fig. 9	Level Lines for Ordinary Minimization Problem	33
Fig. 10	Plots of u , δu^ϕ , δu^Ω and $\delta u'$ as Function of x	35
Fig. 11	Unfiltered Final Distributions, $L=14$, $\ell(x)=0$	40
Fig. 12	Filtered Final Distributions, $L=14$, $\ell(x)=0$	41
Fig. 13	Initial and Final Distributions, $L=14$	42
Fig. 14	Initial and Final Distributions for Different Assumed Initial Distributions, $L=6$, $\ell(x)=0$	44
Fig. 15	Initial and Final Distributions for Different Assumed Initial Distributions, $L=7$, $\ell(x)=0$	45
Fig. 16	Initial and Final Distributions for Different Assumed Initial Distributions, $L=9.9$, $\ell(x)=0$	46
Fig. 17	Initial and Final Distributions for Different Assumed Initial Distributions, $L=17$, $\ell(x)=0$	47
Fig. 18	Initial and Final Distributions for Different Assumed Initial Distributions, $L=18$, $\ell(x)=0$	48
Fig. 18a	Optimum Attenuation as a Function of the Total Length of a Line	49
Fig. 19	Optimal $r(x)$ and $c(x)$ Distribution for $L=14$, $\ell(x)=0$	50

Fig. 20	Optimal $r(x)$ and $c(x)$ Distribution for $L=14, \ell(x)=.05$	51
Fig. 21	Optimal $r(x)$ and $c(x)$ Distribution for $L=14, \ell(x)=0.1$	52
Fig. A-1	Block Diagram for Hybrid Computer	57
Fig. A-2	Logic Patching of AC and AC Counter Output	62
Fig. A-3	Monostable Outputs and Control	62
Fig. A-4	Timing Chart for Integration	62
Fig. B-1	Distributions $\delta u_M(x), \delta u(0), \delta u_m(x), \delta u_p,$ $\delta u_{pa}, \delta u_n, \delta u_{na}$	66

A B S T R A C T

A technique is developed for the synthesis and design of a distributed parameter system guiding waves from one point in space to another. The parameter distributions are assumed to be unrestricted except for the upper and lower bounds resulting from the imposition of physical realizability. The problem is similar to the "sensitivity" problem encountered in the optimal control of the systems. An improved version of the First Order Gradient Technique is used to obtain the optimal distributions of the parameters. The First Order Gradient Technique is sensitive to the form of the arbitrary distributions assumed at the start of the iterations. This technique has serious convergence problems associated with it. The problem is particularly severe and is encountered in the "singular" optimal control problems. The algorithm devised here improves the First Order Gradient Technique so that it becomes less sensitive to the initial assumed distributions and virtually eliminates the convergence problems generated because of the bounds on the parameter distributions.

A transmission line with distributed series r - shunt c is a particular case of the distributed parameter system. The optimal design of a feedback

network, for a phase-shift oscilltor, employing thin film circuit is a successful example of the application of the Improved Gradient Technique. These distributions have been obtained by the use of a Hybrid Computer.

INTRODUCTION

The design of a feedback network for a phase shift oscillator has been a topic of a number of studies. As shown in a very basic block diagram (Fig. 1), the frequency sensitive feedback circuit has to provide a proper gain and phase shift relationship so that the system may oscillate.

Assuming that the amplifier has phase-shift of 180 degrees, the feedback circuit has to provide an additional 180 degree phase shift in order to get the conditions for oscillation. One other usual requirement of the feedback circuit is that the attenuation during the transmission of the signal should be minimum, since the total gain around the loop should be unity. This lowers the gain requirements of the amplifier.

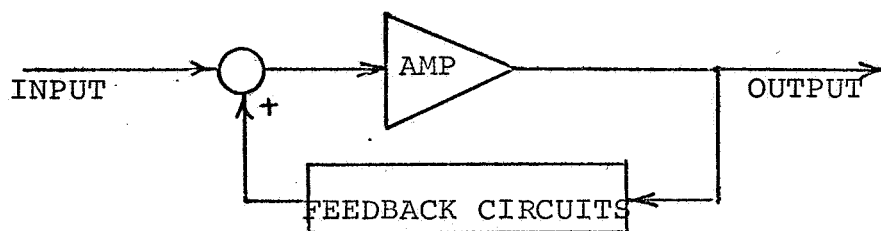


Fig. 1

A three section lumped parameter uniform resistance capacitance network producing 180 degree phase-shift gives¹ an attenuation of 29. Johnson² has shown that the circuit in Fig. 2 gives, in the limit, the attenuation of 8 as K tends to infinity. He also showed that a uniformly distributed-series r , parallel c -network would produce an attenuation of 11.6. Increasing the number of sections in the lumped para-

meter circuit helps to reduce the attenuation. A limiting

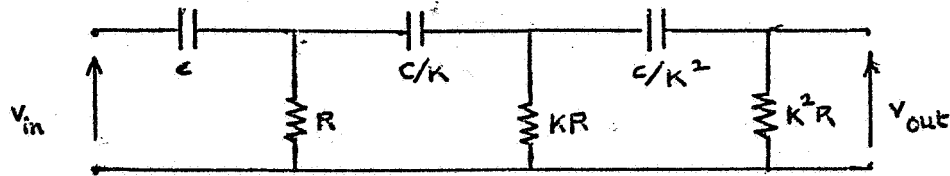


Fig. 2

case is obviously a distributed rc transmission line. Edson³ finds that "unfortunately, the analysis of multiple-section lumped networks is exceedingly complicated and tedious It is found that useful inferences may be drawn from the limiting case in which the number of sections becomes infinite and the network becomes a smoothly tapered transmission line". He assumed the exponential variation of the parameters corresponding to

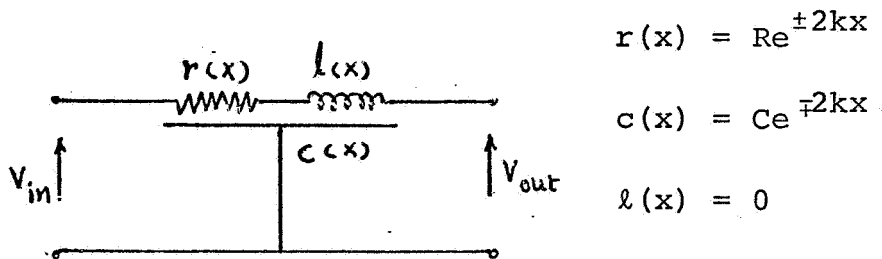


Fig. 3

Edson obtained the curves for attenuation at 180 degree phase-shift as a function of parameters R, C , taper k and line length L . It can be easily shown that as k approaches infinity the attenuation approaches unity.

With the advent of thin-film circuits the distributed parameter RC line ceases to be just a limiting case of n lumped RC circuits. Thin film circuits are replacing the

lumped components due to the requirements of (1) microminimization and (2) modular construction.

A number of materials have been used for thin film resistors, such as vacuum deposited nichrome, sputtered tantalum, vacuum deposited metal oxides, etc. Thin-film capacitors are fabricated by evaporating a high dielectric material onto a resistive path and covering it with another layer of conductive film. The dielectric layer may be formed by oxidizing the resistance layer. These techniques enable us to realize a wide range of resistance and capacitance values per unit length. This can be achieved by controlling the physical dimensions of the films.

Since distributed networks are used extensively in microcircuitry and there is a definite possibility of shaping the distributions so as to optimize the performance of the system in which they are used, the need for developing a technique for such synthesis is apparent. So far, in the field of feedback oscillator circuits, the trend has been to assume certain form of distributions, such as exponentials, and analyze the circuit.

The present study keeps the form of the distributions completely free except for the upper and lower bounds resulting from the physical realizability and tries to obtain optimum distributions of parameters which optimize the specified criterion, viz. minimum attenuation at 180 degree phase shift.

MATHEMATICAL FORMULATION OF THE OSCILLATOR PROBLEM

The general statement of the problem is as follows:

Find the distributions $r(x)$ and $c(x)$ [with reference to Fig. 3] such that at a given frequency (1) there exists desired phase shift between the input and output, and (2) the attenuation is minimum. The distributed inductance $\ell(x)$ is considered to be a non-controllable quantity.

For lumped parameter circuits the system equations governing voltage and current relationships are differential equations in time with constant coefficients. For a distributed parameter system they become partial differential equations in time and space with coefficients being functions of space, [referring to Fig. 4].

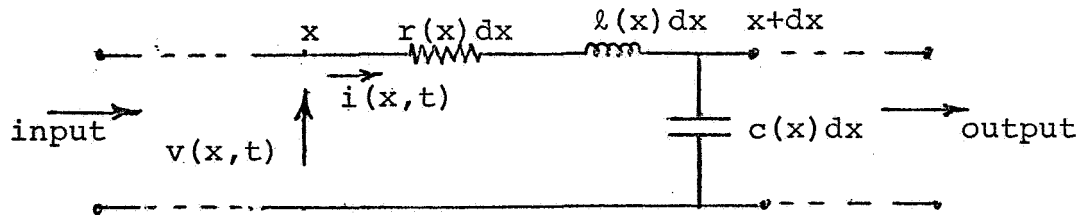


Fig. 4

$$\frac{\partial}{\partial x} [v(x,t)] + r(x)i(x,t) + \ell(x) \frac{\partial}{\partial t} [i(x,t)] = 0 \quad (2.1)$$

$$\frac{\partial}{\partial x} [i(x,t)] + c(x) \frac{\partial}{\partial t} [v(x,t)] = 0 \quad (2.2)$$

The driving function $v(0,t)$ is assumed to be a co-sinusoidal input at a frequency ω . Linearity and the time invariance of parameters r , c , and ℓ assure the presence of only one fre-

quency ' ω '. Thus we can assume a steady state solution of the form

$$v(x,t) = \alpha(x) \cos(\omega t + \theta) = V_1(x) \cos \omega t + V_2(x) \sin \omega t \quad (2.3)$$

$$i(x,t) = I_1(x) \cos \omega t + I_2(x) \sin \omega t \quad (2.4)$$

where $\theta = \tan^{-1} \frac{V_2(x)}{V_1(x)}$ specifies the phase angle of the voltage as a function of x , and $\alpha(x) = (V_1^2(x) + V_2^2(x))^{1/2}$ gives the amplitude of the voltage along the line.

Substituting this solution into the equation (2.1) and (2.2) we obtain time independent state equations

$$\begin{aligned} \frac{d}{dx} V_1(x) &= -r(x)I_1(x,\omega) - \omega l(x)I_2(x,\omega) = f_1 \\ \frac{d}{dx} V_2(x) &= -r(x)I_2(x,\omega) + \omega l(x)I_1(x,\omega) = f_2 \\ \frac{d}{dx} I_1(x) &= -\omega c(x)V_2(x) = f_3 \\ \frac{d}{dx} I_2(x) &= \omega c(x)V_1(x) = f_4 \end{aligned} \quad (2.5)$$

At this stage we will make two assumptions, (1) output impedance of the amplifier [source impedance at the input of the line] is zero and (2) input impedance of the amplifier [load impedance on the line] is infinite.

Without any loss of generality the input conditions of the line could be specified as

$$\begin{aligned} V_1(0) &= a, \quad a > 0 \\ V_2(0) &= 0 \end{aligned} \quad (2.6)$$

The open circuit at the output end of the line implies

$$I_1(L) = I_2(L) = 0. \quad (2.7)$$

The two conditions given above get slightly modified for a non-zero amplifier output impedance and finite amplifier input impedance.

The 180 degree phase shift requirement is translated as

$$\begin{aligned} V_2(L) &= 0 \\ V_1(L) &< 0. \end{aligned} \quad (2.8)$$

Equation (2.8) assures a phase shift of $\pi, 3\pi, 5\pi, \dots$. Since the minimum attenuation is the same as the maximum gain we need to maximize ϕ ,

$$\phi = \left| \frac{\alpha(L)}{\alpha(0)} \right| \quad (2.9)$$

with $\alpha(x)$ as defined in (2.3).

Since $V_2(0) = V_2(L) = 0$ and $V_1(0) = a$, the criterion becomes

$$\max \phi = \max |V_1(L)|. \quad (2.10)$$

The general form of the solution will be as shown in Fig. 5.

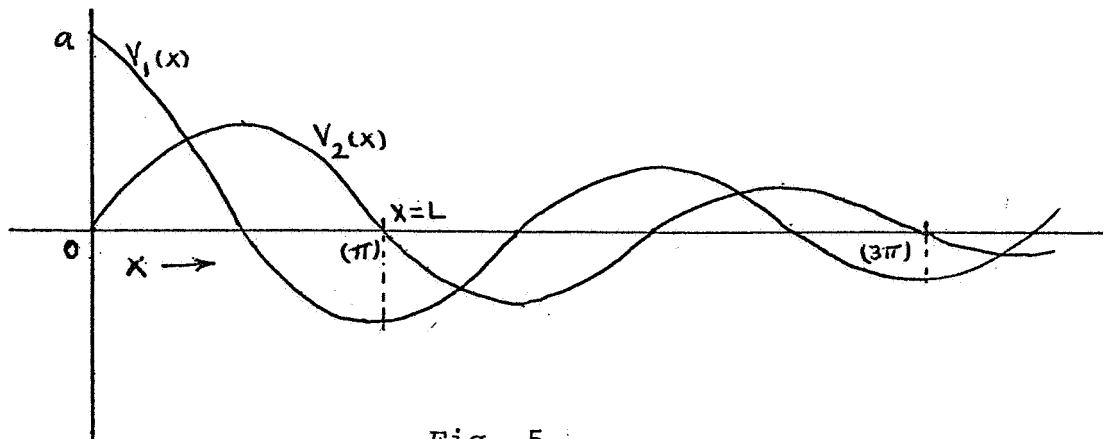


Fig. 5

The attenuation increases as the signal travels along the line. Thus there is no possibility of attenuation at phase shift of 3π , 5π ... being smaller than that at π . We can safely restrict our considerations to the phase shift of π , or the first zero of $V_2(x)$.

The boundary conditions (2.6) and (2.7) require solving a two point boundary value problem, since the voltage is specified at one end and the current at the other end. It is possible to avoid mixed boundary conditions by specifying the voltage at $x=L$.

If the conditions are specified as $V_1(L) = a$, $V_2(L) = 0$, from (2.9) it is apparent that, with

$$V_2(0) = I_1(L) = I_2(L) = 0 ,$$

$$\max \phi = \min |V_1(0)| .$$

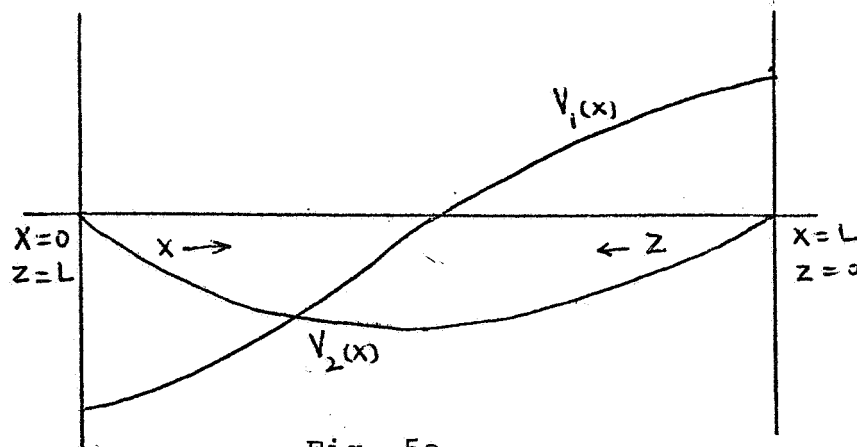


Fig. 5a

Fig. 5a gives a general form of the solution. For $a > 0$, $V_1(x=0) < 0$. This implies that

$$\max \phi = \max V_1(0) . \quad (2.11)$$

Now we can define the problem using control system terminology. Define the 4-vector

$$y^t = [V_1, V_2, I_1, I_2].$$

The system equations are given by

$$\frac{d}{dx} y(x) = A(u(x)) y(x). \quad (2.12)$$

The matrix $A(u(x))$ is defined by (2.5) and $u(x)$ is a two-vector defined by $u^t(x) = [r(x), c(x)]$. The inductance $\ell(x)$ is assumed to be a non-controllable parameter. The endpoint boundary conditions are

$$y^t(x=L) = [a, 0, 0, 0] \quad (2.13)$$

with L fixed; and the rigid constraint Ω is given by

$$\Omega[y(x=0)] = y_2(0) = 0. \quad (2.14)$$

Our task is to obtain r and c distributions that maximize ϕ , where

$$\phi = y_1(0). \quad (2.15)$$

We also assume that the limitations in fabrication require that the values of resistance and capacitance per unit length be within finite upper and lower bounds. This gives rise to the inequality constraints on the control variables r and c .

$$\begin{aligned} r_m &\leq r(x) \leq r_M \\ c_m &\leq c(x) \leq c_M \end{aligned} \quad (2.16)$$

METHODS OF SOLUTION

A. Hamilton-Jacobi Equations via Dynamic Programming:

This was tried as a possible approach. It is presented here to give some idea about the complexity involved in the numerical solutions of the two point boundary value problem one may face in using techniques that lead to a set of necessary conditions for optimality.

Let us define a new independent variable

$$z = L - x. \quad (3.1)$$

The state equation which was the same as (2.12) now becomes

$$\frac{d}{dz} y(z) = -A(u(z))y(z) \quad (3.2)$$

and the end point conditions specified in (2.11) now become initial conditions

$$y(z=0) = y_L = y^0. \quad (3.3)$$

With the criterion function as $\phi = \phi(y(z=L))$ we have a Mayer formulation of the variational problem. Bellman and Dreyfus⁷ have used a heuristic approach that is very revealing. The optimal payoff function as designated by J , is an implicit function of the initial state $y^0 = y(z_0)$ and the length of the process

$$s = L - z_0. \quad (3.4)$$

The optimal payoff J is defined by

$$J = J(y^0, s) = \max_u [\phi(\bar{y}(L))] \quad (3.5)$$

The optimal vector $u^*(z)$ and the optimal state vector $y^*(z)$ all depend on y° and s . Consider a distance Δz along the optimal trajectory as shown in Fig. 6.

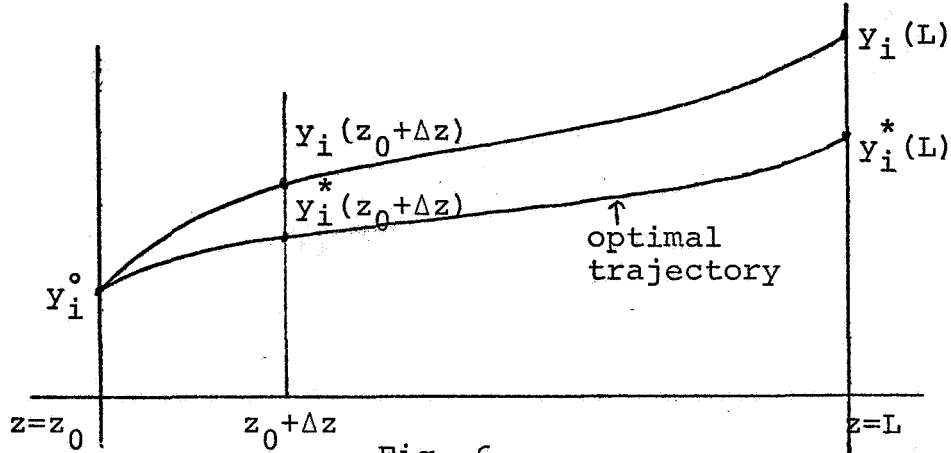


Fig. 6

The trajectory for nonoptimal u is also shown. Along the optimum path,

$$J(y^\circ, s) = J(y^*(z_0 + \Delta z), s - \Delta z), \quad (3.6)$$

since we will end up at $y^*(L)$ along the optimum path, no matter where we start from.

If we take an arbitrary u from z_0 to $z_0 + \Delta z$ and an optimal $u = u^*$ from $z_0 + \Delta z$ to L then the payoff function will be $J(y(z_0 + \Delta z), s - \Delta z)$.

Thus,

$$J(y^\circ, s) = \max_u [J(y(z_0 + \Delta z), s - \Delta z)]. \quad (3.7)$$

Expanding the right side in a Taylor series and neglecting second and higher order terms,

$$\begin{aligned} J(y^\circ, s) &= \max_u [J(y^\circ, s) - \frac{\partial J}{\partial s} \Delta z + \sum_i \frac{\partial J}{\partial y_i} \Delta y_i] \\ &= \max_u [J(y^\circ, s) - \frac{\partial J}{\partial s} \Delta z + \sum_i \frac{\partial J}{\partial y_i} \frac{d}{dz} y_i \Big|_{(y^\circ, z_0)} \Delta z], \end{aligned} \quad (3.8)$$

since

$$\Delta y_j = \frac{d}{dz} y_j \Delta z. \quad (3.9)$$

The term on the right side which depends on u is $\frac{d}{dz} y_i(z_0)$. Hence we take all the other terms outside of the bracket and divide by Δz . Taking the limit as $\Delta z \rightarrow 0$ we obtain

$$\frac{\partial J}{\partial s} = \max_{u(z_0)} \left[\sum_j \frac{\partial J}{\partial y_j} \frac{d}{dz} y_j \right]_{(y^0, z_0)} \quad (3.10)$$

In (3.6) the maximization was to be carried out over the interval z_0 to $z_0 + \Delta z$. With $\Delta z \rightarrow 0$ the control vector becomes just $u(z_0)$.

The partial differential equation (3.10) is true for any z along the trajectory and the corresponding duration s . Thus,

$$\frac{\partial J}{\partial s} = \max_{u(z)} \left[\sum_i \frac{\partial J}{\partial y_j} \frac{d}{dz} y_j \right]_{(y, z)}, \quad (3.11)$$

This is a Hamilton-Jacobi equation.

For our case y is a 4-vector and u is a 2-vector, $u^t = [r(z), c(z)]$. We can incorporate the constraint (2.14) into the criterion function by means of a Lagrange multiplier μ and write a new payoff function,

$$\phi = y_1(z=L) + \mu y_2(z=L). \quad (3.12)$$

In order to solve (3.7) numerically, we have to discretize it in z . As a first step we have to obtain solution of $u(z_0)$ and then for $u(z_0 + \Delta z)$, assuming $u(z_0)$ to be constant from z_0 to $z_0 + \Delta z$, and so on. This becomes a problem of grid formation⁸ in five dimensional space $[y_1, y_2, y_3, y_4, s]$. The undetermined Lagrange multiplier μ is an unknown quantity that has to be determined by trial and error.

No attempt was made to obtain the numerical solutions using this approach.

B. Pontryagin's Maximum Principle:

The maximum Principle will yield a set of necessary conditions that the optimal control u^* has to satisfy if such control exists, and if it optimizes the criterion function.

Given the system equations (restatement of (3.2) and (3.3),

$$\frac{d}{dz} y(z) = -A(u(z))y(z), \quad (3.13)$$

with boundary conditions

$$y(z=0) = y^0 = [a, 0, 0, 0], \quad (3.14)$$

the Maximum Principle states that in order that the trajectory $y(z)$ be optimal in the sense that the criterion function is maximized it is necessary that one can find functions $\lambda(z)$, defined as the adjoint variables, satisfying the following properties,

i) the λ_j satisfy the differential equations,

$$\frac{d}{dz} \lambda_j + \sum_{i=1}^4 \left[\frac{d}{dy_j} f_i(y^*(z), u^*(z)) \right] \lambda_i(z) = 0 \quad (3.15)$$

ii) letting

$$E(u) = \sum_{i=1}^4 f_i(y(z), u) \lambda_i(z)$$

where u is arbitrary, nonoptimal 2-vector, then

$$E(u^*(z)) \geq E(u(z)) \quad \text{for all } z \text{ and admissible } u. \quad (3.16)$$

At this stage we may introduce the Hamiltonian,

$$H(y, u, \lambda) = \sum_{j=1}^4 \lambda_j f_j(y, u) \quad (3.17)$$

Thus, in terms of the Hamiltonian the Maximum Principle states that, given

$$\frac{d}{dz} y = H_{\lambda}^t \quad (3.18)$$

$$\frac{d}{dz} \lambda = -H_y^t$$

in order that $y(z)$ maximizes the criterion function it is necessary that

$$H(y(z), u(z), \lambda(z)) \geq H(y(z), \bar{u}, \lambda(z)) \quad (3.19)$$

where \bar{u} is any constant admissible control vector.

Also the adjoint variables have to satisfy the transversality condition at $z=L$

$$\lambda_j(L) = \frac{\partial \phi}{\partial y_j} - \mu \frac{\partial \Omega}{\partial y_j}, \quad (3.20)$$

where ϕ is the criterion function, Ω is the boundary constraint and μ is the undetermined Lagrange Multiplier.

Coming back to our system of equations,

$$\frac{d}{dz} y = -Ay$$

$$\frac{d}{dz} \lambda = A^t \lambda$$

$$H = -\lambda^t A y \quad (3.21)$$

$$y^t(0) = [a, 0, 0, 0]$$

$$\lambda^t(L) = [1, -\mu, 0, 0].$$

With the inequality constraint (2.16) on the control, (3.19) gets transformed into

$$\begin{aligned}
 \frac{\partial H}{\partial u^*_i} &> 0 && \text{if } u^*_i = u_{imax} \\
 \frac{\partial H}{\partial u^*_i} &= 0 && \text{if } u_{imax} > u^*_i > u_{imin} \\
 \frac{\partial H}{\partial u^*_i} &< 0 && \text{if } u^*_i = u_{imin}
 \end{aligned} \tag{3.22}$$

and the constraining equation

$$\Omega[y(L)] = 0 \tag{3.23}$$

Equations (3.21) to (3.23) comprise a self-sufficient set of equations, which, when solved will yield the optimal control $u^*(z)$. This is a two point boundary value problem.

Since $A(u(z))$ is linear in u , H turns out to be linear in u .

$$\begin{aligned}
 H_r &= (\lambda_1 y_3 + \lambda_2 y_4) , \\
 H_c &= \omega(\lambda_3 y_2 - \lambda_4 y_1) .
 \end{aligned} \tag{3.24}$$

Thus (3.19) suggests a bang bang control. In other words, we are tempted to believe that u_i will always be at either boundary and will switch whenever H_{u_i} changes sign.

Assuming such a bang bang form of the control, a combination of iterations and scanning (for μ) was used to find a solution. The iterations did not converge. The solutions obtained by gradient technique which is described later, indicate that the assumption regarding the form of the control was erroneous. It does not turn out to be a bang bang control.

(One way of circumventing the problem of linearity is by adding to the criterion function a penalizing functional⁷ that is nonlinear in control).

As Johnson and Gibson⁴ have pointed out, it is characteristic of the solutions to linear optimization problems that the switching function H_{u_i} sometimes becomes identically zero over some finite interval of 'z'. Since, during this interval, H does not depend upon u explicitly, the usual procedure of selecting u^* so as to maximize H breaks down. These linear optimization problems where H_{u_i} becomes identically zero over finite interval have been referred to as "singular". It has been shown that the optimal control may actually consist of intervals of variable control effort (called "singular switching curves") combined with intervals of limiting control.

Thus, there seems to be a distinct possibility of the optimal control being a limiting control with singular curves rather than a bang bang control with switching points.

C. Gradient Technique⁵:

The approaches described so far are based on obtaining a set of necessary conditions for the optimality and then trying to get solutions to this set of equations. The necessary condition may be a partial differential equation as in the case of the Hamilton-Jacobi equation or a set of differential equations with mixed boundary conditions as for the Maximum Principle. The approximate solutions to these equations may not lie in the neighborhood of the desired solution and thus

may not yield any useful information.

The alternate approach is to seek stepwise gradual improvement in the criterion function. The method is known as "Gradient Technique", "Hill Climbing Technique", or "relaxation method". Here one assumes an arbitrary non-optimal solution and seeks a stepwise improvement in the direction of the optimum. Thus, the new solution generated at every step of the iteration is an improvement over the previous one and the process hopefully converges.

Specifically, we seek to obtain a functional relationship between variations in the criterion function and variations in the control vector u . This defines the desired variation in u in order to achieve improvement in ϕ and yields a self-sufficient iteration procedure.

Referring to the set of equations (2.5) we have

$$\frac{d}{dx} y_i = f_i(y, u).$$

Consider a small perturbation in the control variable u . With new control as $u + \delta u$ and resulting trajectories as $y + \delta y$ the resulting first order variational equations are given by,

$$\frac{d}{dx} (\delta y_i) = \sum_j \frac{\partial f_i}{\partial y_j} \delta y_j + \sum_k \frac{\partial f_i}{\partial u_k} \delta u_k, \quad (3.22)$$

i.e.

$$\frac{d}{dx} (\delta y) = A \delta y + A_u \delta u; \quad (3.23)$$

A_u denotes the partial differentiation of the 'A' matrix with reference to the subscripted variable.

Define a set of adjoint variables by the differential equations,

$$\frac{d}{dx} \lambda_i(x) = - \sum_j \frac{\partial f_j}{\partial y_i} \lambda_j(x). \quad (3.24)$$

The system equations are linear in y . Under such circumstances the adjoint equations will always be reduced to a form

$$\frac{d}{dx} \lambda(x) = - A^t(u(x)) \lambda(x). \quad (3.25)$$

Before illustrating the function of the adjoint variables, we can discuss their form.

From (2.12) and (3.25) one can derive the relationship

$$\frac{d}{dx} (\lambda^t y) = 0 \quad (3.26)$$

i.e. the inner product of λ and y remains constant for all x .

This implies,

$$\lambda^t(0)y(0) = \lambda^t(L)y(L) \quad (3.27)$$

In terms of fundamental matrices $\Phi(x)$ and $\Psi(x)$, where

$$y(x) = \Phi(x)y(0) \quad (3.28)$$

$$\lambda(x) = \Psi(x)\lambda(0) \quad (3.29)$$

the relationship (3.27) yields

$$\Psi(x) = \Phi^{-1}(x) \quad (3.30)$$

Hence, if the solutions to (2.5) are known in terms of the fundamental matrix Φ , the solutions for the adjoint equations can be obtained as

$$\lambda(x) = \Phi^{-1}(x)\lambda(0)$$

without solving (3.25).

Also for certain forms of matrix A there exists a simple linear transformation of the type

$$\gamma(x) = B\lambda(x) \quad (3.31)$$

where B is a nonsingular constant matrix, such that with this transformation the adjoint equations

$$\frac{d}{dx} \lambda = -A^t \lambda$$

become

$$\frac{d}{dx} \gamma = -A\gamma$$

With the change of variable as

$$z = L - x \quad (3.32)$$

the equations reduce to

$$\frac{d}{dz} \gamma(z) = A(z)\gamma(z) \quad (3.33)$$

This equation has a form identical to (2.12). Thus the transformed adjoint variables $\gamma(z)$ are solutions of the system equations with the reversal of the space variable. The 'A' matrix, as defined by (2.5) and (2.12), possesses the above properties. We will consider the corresponding B matrix and the significance of the property mentioned above later, when we will get to the stage of obtaining the numerical solutions to the system and the adjoint equations.

Multiplying (3.22) by λ_i , (3.24) by δy_i , adding them together and performing summation over i , we get

$$\frac{d}{dx} [\sum_i \lambda_i \delta y_i] = \sum_i \sum_k \lambda_i \frac{\delta f_i}{\delta u_k} \delta u_k \quad (3.34)$$

Define

$$H = \langle \lambda, f \rangle = \sum_i \lambda_i f_i .$$

Then, integrating (3.34) from $x=0$ to $x=L$,

$$[\sum_i \lambda_i \delta y_i]_0^L = \int_0^L H_u \delta u dx . \quad (3.35)$$

Since $y(L)$ is completely specified by (2.13),

$$\delta y(L) = 0 .$$

Here we can define λ^ϕ and λ^Ω as the adjoint system variables satisfying (3.25) subject to boundary conditions,

$$\lambda_i^\phi(0) = - \frac{\partial \phi}{\partial [y_i(0)]} \quad (3.36)$$

and

$$\lambda_i^\Omega(0) = - \frac{\partial \Omega}{\partial [y_i(0)]} \quad (3.37)$$

respectively.

When ϕ is given by (2.15), equation (3.36) becomes

$$\lambda^\phi(0) = [-1, 0, 0, 0]^t .$$

Similarly, when Ω is given by (2.14), equation (3.37) becomes

$$\lambda^\Omega(0) = [0, -1, 0, 0]^t .$$

Now we can define

$$H^\phi = \langle \lambda^\phi, f \rangle$$

$$H^\Omega = \langle \lambda^\Omega, f \rangle .$$

Substituting (3.36) into (3.35), we obtain

$$\delta \phi = \int_0^L H_u^\phi \delta u dx . \quad (3.38)$$

Similarly, substituting (3.37) into (3.35), we obtain

$$\delta\Omega = \int_0^L H_u^\Omega \delta u dx \quad . \quad (3.39)$$

Equations (3.38) and (3.39) give the functional relationship between variation in control, and change in criterion function and constraint in response to it.

The initial arbitrary nonoptimal choice of u or the subsequent estimates of u during iteration process may not exactly satisfy the constraint $\Omega=0$. Therefore, at every stage there are two variations required.

- (i) Change $d\phi$ in order to improve the criterion function.
- (ii) Change $d\Omega=-\Omega$ in order to satisfy the constraining equation.

Noting that δu is a function of 'x' one realizes that (3.37) and (3.38) can have infinity of solutions. Hence we stipulate an arbitrary criterion function $1/2 \int_0^L \delta u^t W \delta u dx$ which has to be minimized while satisfying (3.38) and (3.39). This (i) eliminates the singular problem since the criterion is quadratic in control, and (ii) keeps the variation δu to a minimum (in Euclidian norm sense). This is desirable since the derivations are based on a small perturbation.

We have to find δu that minimizes the composite criterion function

$$\psi = 1/2 \int_0^L \delta u^t W \delta u dx + v^\phi [d\phi - \int_0^L H_u^\phi \delta u dx] + v^\Omega [d\Omega - \int_0^L H_u^\Omega \delta u dx], \quad (3.40)$$

where v^ϕ and v^Ω are undetermined Lagrange multipliers, to be chosen so as to satisfy (3.38) and (3.39).

Euler Lagrange equations give

$$\delta u = W^{-1} [v^\phi H_u^{\phi t} + v^\Omega H_u^{\Omega t}] \quad (3.41)$$

Defining

$$\delta u^\phi = W^{-1} H_u^{\phi t} \quad , \quad (3.41a)$$

$$\delta u^\Omega = W^{-1} H_u^{\Omega t} \quad , \quad (3.41b)$$

with

$$H_u^t = \begin{bmatrix} -(\lambda_1 y_3 + \lambda_2 y_4) \\ \omega(\lambda_4 y_1 - \lambda_3 y_2) \end{bmatrix}$$

$$\delta u = v^\phi \delta u^\phi + v^\Omega \delta u^\Omega \quad (3.42)$$

Substituting for δu in (3.38) and (3.39) we get

$$d\phi = v^\phi \int H_u^\phi \delta u^\phi dx + v^\Omega \int H_u^\phi \delta u^\Omega dx \quad (3.43)$$

$$d\Omega = v^\phi \int H_u^\Omega \delta u^\phi dx + v^\Omega \int H_u^\Omega \delta u^\Omega dx \quad (3.44)$$

The iteration algorithm is fairly straightforward and proceeds as follows.

1) Assume a nominal control u . Solve system equations (2.12) with boundary conditions (2.13).

2) Solve the adjoint equations (3.25): (i) with boundary condition (3.36) to obtain $\lambda^\phi(x)$ and (ii) with boundary condition (3.37) to obtain $\lambda^\Omega(x)$.

3) Evaluate δu^ϕ and δu^Ω from (3.41a) and (3.41b), with W^{-1} given.*

4) Solve (3.43) and (3.44) for v^ϕ and v^Ω .**

5) Evaluate δu from (3.42).

6) Add δu to u & obtain revised estimate for control as

* W^{-1} is chosen based on knowledge of the system, & could be made equal to the unity matrix.

** $d\phi$ & $d\Omega$ must be chosen beforehand. $d\phi$ is chosen for convergence & $d\Omega$ to satisfy the constraint equation (2.14).

$u + \delta u$. The inequality constraints on u are taken into account by truncating $(u + \delta u)$ at u_M or u_m wherever it crosses the bounds. The validity of truncation in connection with convergence can be proven for ϕ or Ω correction separately. (Refer Appendix B) Now branch back to the start of the loop for the next iteration cycle.

The variation $u(x)$ has two components $\delta u^\phi(x)$ and $\delta u^\Omega(x)$ and their proportion in forming δu is decided by two scalar coefficients v^ϕ and v^Ω . But suppose we have only one criterion function ϕ and no constraint. Then (3.43) will be transformed into

$$\delta u = v^\phi \delta u^\phi$$

and v^ϕ can be obtained from (3.43) by setting $\delta u^\Omega = 0$.

There are two aspects of the form of δu which we can control:

1. The shape of the variation in $u(x)$; δu^ϕ and δu^Ω .
2. The amount of variation or the step size which is constant for all x ; v^ϕ and v^Ω .

Equation (3.41) gives the shape of $\delta u(x)$ as

$$\delta u^\phi(x) = W^{-1} H_u^{\phi t}$$

and v^ϕ is the step size.

Thus instead of viewing $W(x)$ as a weighting factor for an arbitrary criterion function we can choose $W^{-1}(x)$ as a shaping factor and obtain the corresponding step size directly from (3.38).

Let us consider an analogy from the field of calculus. Let the criterion function ϕ which has to be minimized be

a function of two independent variables x_1 and x_2 .

$$\phi = \phi(x_1, x_2) . \quad (3.45)$$

Hence the first order variational equation is

$$d\phi = \phi_{x_1} \delta x_1 + \phi_{x_2} \delta x_2 , \quad (3.46)$$

where ϕ_{x_1} and ϕ_{x_2} are partial differentials of ϕ with respect to x_1 and x_2 respectively.

For a given $d\phi$ we can find non-unique values of δx_1 and δx_2 . An additional constraint that removes the non-uniqueness is obtained when we seek a variation $[\delta x_1, \delta x_2]$ that (i) minimizes

$$1/2 ||\delta x||^2 = 1/2 (\delta x_1^2 + \delta x_2^2) , \quad (3.47)$$

and (ii) satisfies (3.46).

The composite criterion function for this accessory minimization problem can be written down (Refer (3.40)).

$$\psi = 1/2 (\delta x_1^2 + \delta x_2^2) + v[d\phi - (\phi_{x_1} \delta x_1 + \phi_{x_2} \delta x_2)]$$

The conditions for stationarity of ψ yield

$$\begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix} = v \begin{bmatrix} \phi_{x_1} \\ \phi_{x_2} \end{bmatrix} \quad (3.48)$$

Substituting this back into (3.46)

$$v = d\phi / (\phi_{x_1}^2 + \phi_{x_2}^2)$$

In equation (3.48) the gradient $\begin{bmatrix} \phi_{x_1} \\ \phi_{x_2} \end{bmatrix}$ gives the direction in

(x_1, x_2) space and v is the step size.

Referring back to the problem of the calculus of variations, if we let $W^{-1} = \text{identity matrix}$, δu^ϕ in equation (3.41) is the 'gradient' of ϕ at any x . We can call this function a 'shape' of the variation that specifies the 'direction' in y space at all ' x '. The constant v^ϕ in (3.42) is comparable to the step size v in the above example.

Thus, the separation of the variation in control as a 'shape factor' δu^ϕ and a step size v^ϕ is comparable to the 'gradient' and a step size.

ALGORITHMS AND PROGRAMS

The iterative solutions are obtained on the Hybrid Computer Unit using the improved Gradient Technique. Appendix A describes the features and certain operations of the Hybrid Unit. The analog computer is used exclusively for solving the differential equations. The digital computer supplies the continuously varying coefficients. The synchronous operation of the analog and digital computer units yields the solutions to the differential equations. The solutions to the various system equations are stored into the memory of the DC and are subsequently operated upon to obtain the desired variation in the control variables. The entire operation is under complete program control of the DC.

Analog Patching:

We need to solve three sets of system equations on the analog computer.

The system proper is described by (from (3.1))

$$\begin{aligned}
 \frac{d}{dz} y_1(z) &= r(z)y_3(z) + \omega l(z)y_4(z), \\
 \frac{d}{dz} y_2(z) &= r(z)y_4(z) - \omega l(z)y_3(z), \\
 \frac{d}{dz} y_3(z) &= c(z)\omega y_2(z), \\
 \frac{d}{dz} y_4(z) &= -c(z)\omega y_1(z).
 \end{aligned}
 \tag{4.1}$$

and

$$y^t(z=0) = [a, 0, 0, 0] \tag{4.2}$$

As shown in Fig. 6, 'x' is a forward and 'z' is a backward direction of integration. The independent variable for an analog computer is time 't'. The AC always integrates forward in time 't'. By setting $t=z$ the equations (4.1) are integrated backwards in space.

The two adjoint systems have identical differential equations (from (3.23)),

$$\begin{aligned}\frac{d}{dx} \lambda_1(x) &= -c(x)\omega\lambda_4(x), \\ \frac{d}{dx} \lambda_2(x) &= c(x)\omega\lambda_3(x), \\ \frac{d}{dx} \lambda_3(x) &= r(x)\lambda_1(x) - \omega\ell(x)\lambda_2(x), \\ \frac{d}{dx} \lambda_4(x) &= r(x)\lambda_2(x) + \omega\ell(x)\lambda_1(x).\end{aligned}\tag{4.3}$$

With the boundary conditions specified by $\lambda(x=0) = \lambda^\phi(0) = [-1, 0, 0, 0]^t$ the solutions of (4.3) give $\lambda^\phi(x)$ and with

$$\lambda(x=0) = \lambda^\Omega(0) = [0, 1, 0, 0]^t\tag{4.5}$$

the solutions of (4.3) give $\lambda^\Omega(x)$.

Equations (4.1) and (4.3) have the same form. With the transformation

$$\begin{aligned}p_1 &= y_1(z) = \lambda_4(x), \\ p_2 &= y_2(z) = \lambda_3(x), \\ p_3 &= y_3(z) = \lambda_2(x), \\ p_4 &= y_4(z) = \lambda_1(x).\end{aligned}\tag{4.6}$$

and with the proper choice of space varying coefficients $r(x)$, $c(x)$, and $l(x)$ and the initial conditions, the same set of equations yields solutions for either $y(z)$, $\lambda^\phi(x)$ or $\lambda^\Omega(x)$.

The 'B' matrix referred to in (3.31) turns out to be

$$B = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} .$$

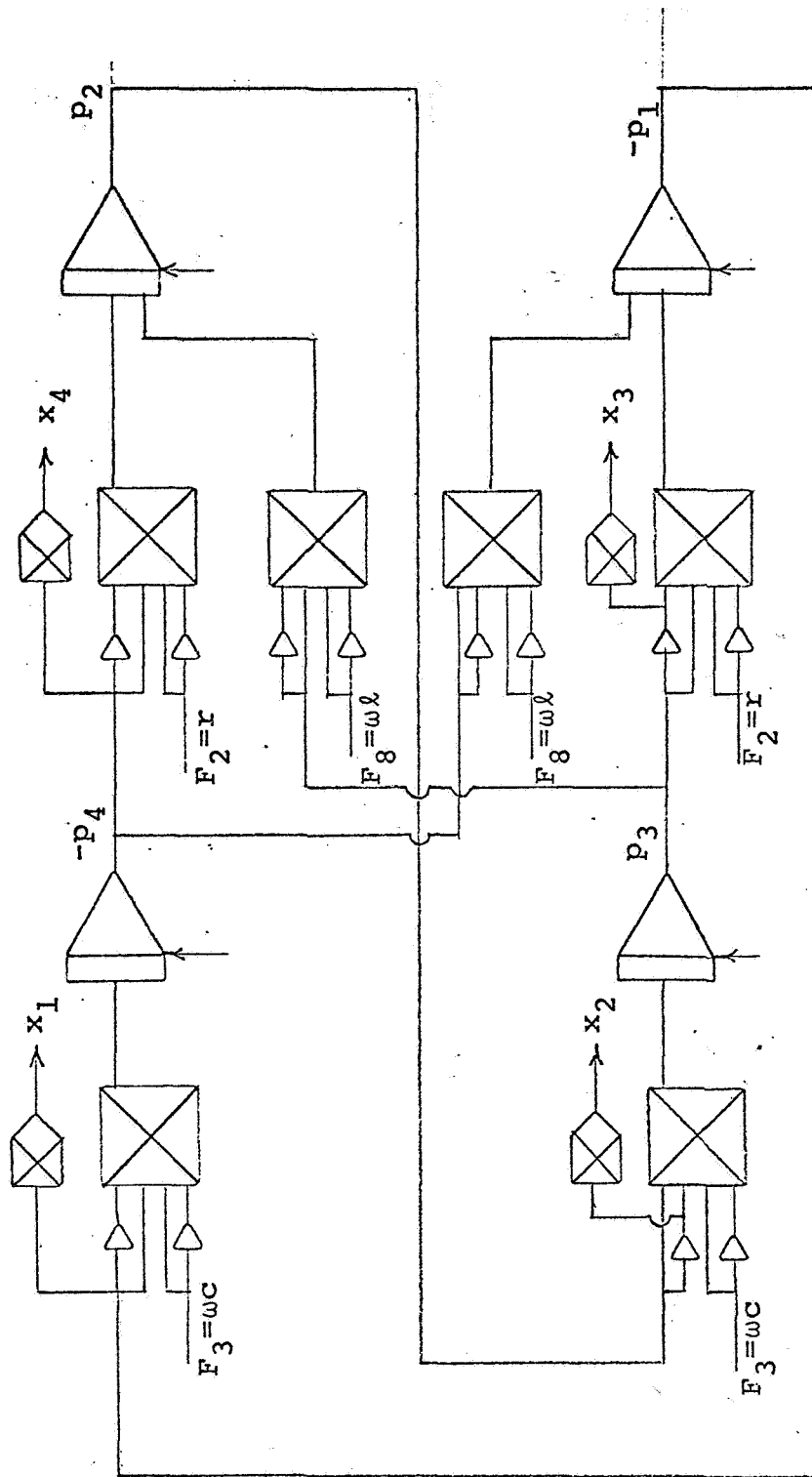
The analog computer patching is given in Figure 7.

A. Algorithm for Improve First Order Gradient Technique:

The flow chart in Fig. 8 describes the hybrid program for the first order estimation of the correction by the improved gradient technique. A more elaborate description is given below.

Block 1: Preparatory Steps -- The input/output channels of the DC are reset; the length of integration is specified; the quantum of the x or z interval which results from discretization of the space is calculated. (The functions $r(x)$ and $c(x)$ are approximated by the staircase approximation.) The upper and lower limits on the control variables are specified and the arbitrary initial profile of the control variables is assumed and loaded into the memory.

Block 2: Solving the System Equations on Hybrid Unit -- The DC sets the initial conditions for the integrators of the AC as given by (4.2). The initial values for the integrators can be obtained either by (a) using the DC to set a pot or (b) using DAC output lines. The initial values of the functions $r(z)$ and $c(z)$ are set up on the DAC. The static test may be



$$p_1 = y_1(z) = \lambda_4(x)$$

$$p_2 = y_2(z) = \lambda_3(x)$$

$$p_3 = y_3(z) = \lambda_2(x)$$

$$p_4 = y_4(z) = \lambda_1(x)$$

INTEGRATOR

INVERTOR

MULTIPLIER

TRACK & STORE

FIGURE 7

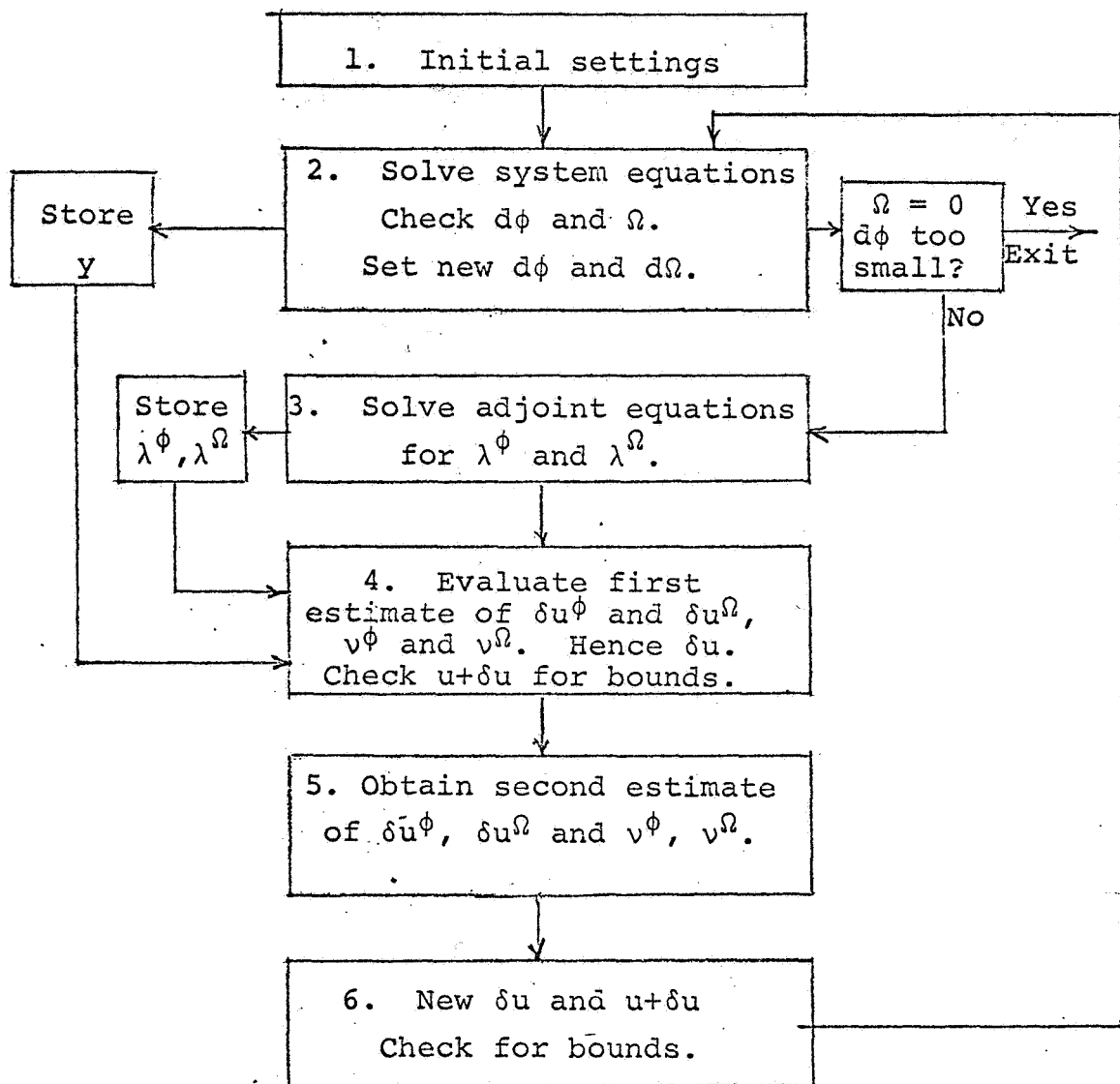


FIGURE 8

carried out at this time to check the initial conditions.

The integration routine then follows. The AC solves the system differential equations. The equations specified by (5.1) are integrated backwards in space. During integration the AC receives from the DC the values of the variable coefficients on the DAC and transmits back the values of the system variables on the ADC.

The solutions obtained by integration are converted to digital form and are stored in the memory of the DC. The values of the criterion function and the residue for the constraining equation are evaluated as,

$$\phi = y_1(z=L)$$

$$\Omega = y_2(z=L).$$

The variation in the criterion function, $d\phi$, is chosen so as to drive $y_1(x=0)$ towards the value of $y_1(L)$ such that the attenuation approaches unity and $d\Omega$ is chosen so that Ω constraint is rigorously satisfied by the next set of distributions, i.e.

$$d\Omega = -\Omega \quad (\text{or } \Omega + d\Omega = 0)$$

Block 3: Solving the Adjoint Equations on the Hybrid Unit -- The operations are identical to the previous block except that (a) the initial conditions are specified by (4.4) and (4.5) for λ^ϕ and λ^Ω respectively, and (b) the adjoint equations (4.3) are integrated forward in space so that $x=t$. Hence the control distributions are $r(x)$ and $c(x)$. The same analog program that is used for the system equations is used

for the adjoint systems with the transformation of variables as given by (4.6).

Block 4: The First Estimate of δu -- The equations (3.41), (3.43), (3.44) and (3.42) yield the estimate of δu . (W is assumed to be an identity matrix.) Adding δu to u one gets an estimate of the new control as $u + \delta u$. However, if u lies close to or is equal to the limiting values the new control $u + \delta u$ may exceed the limits. Under these circumstances $u + \delta u$ is confined to the limiting values wherever it exceeds the limits on the control variables. This amounts to the truncation of δu so that $u + \delta u$ lies within the specified limits (see Fig. 10).

The estimated new control is monitored at this point to check if it exceeds the bounds and truncated if necessary. In the case of the unimproved gradient technique, the program branches back from here to block 2 and starts the new iteration loop.

It is observed that after the control variables reach the limiting values and start getting truncated, the subsequent iterations improve ϕ but cause Ω to diverge instead of converging to zero. It does not pay (in terms of convergence) to let Ω diverge too much. It becomes necessary to set a limit for $|\Omega|$ and monitor it at every iteration.

Whenever Ω diverges and exceeds the limit, only Ω correction is applied during the iteration by assuming $\delta u^{\phi=0}$ in equation (3.42). Thus

$$\delta u = v^{\Omega} \delta u^{\Omega}$$

during such iterations. This procedure drives Ω close to zero without any regard to the value of ϕ . When $|\Omega|$ is driven sufficiently below the limiting value the imposed restriction ' $\delta u^\phi = 0$ ' can be removed and one can seek both ϕ and Ω corrections simultaneously.

So far we have not said anything about what values the elements of weighting matrix W should have. The matrix W has to be positive so as to satisfy the strengthened Legendre necessary condition for the accessory problem. Normally W is chosen to be an identity matrix. In such a case H_u solely determines the shape of δu . H_u is also the 'sensitivity function' or the 'influence function' for the improvement (see equation (3.38) and (3.39)). If we define $u_{\text{opt}}(x)$ as the optimal distribution of scalar $u(x)$, the variation needed to reach the optimal distribution from $u(x)$ will be $(u_{\text{opt}} - u)$. For various values of x the values of $H_u(x)/(u_{\text{opt}}(x) - u(x))$ may turn out to be very different. The control $u(x)$ may be already close to the optimal profile in the most sensitive regions and farther away in the least sensitive regions. The matrix $W(x)$ can be used in such cases as a compensating factor. The matrix

$$W^{-1} = \begin{bmatrix} \left(\frac{r_M - r_m}{L} \right) x + r_m & 0 \\ 0 & c_M - \left(\frac{c_M - c_m}{L} \right) x \end{bmatrix} \quad (4.7)$$

has been found to be helpful in the present case. The choice of W was governed by the sensitivity.

However, once δu starts getting truncated at the boundaries, we face a different type of convergence problem. In order to get a better understanding we will first consider a simpler analogy.

Let $\phi = \phi(x_1, x_2)$ be the cost function of scalars x_1 and x_2 .

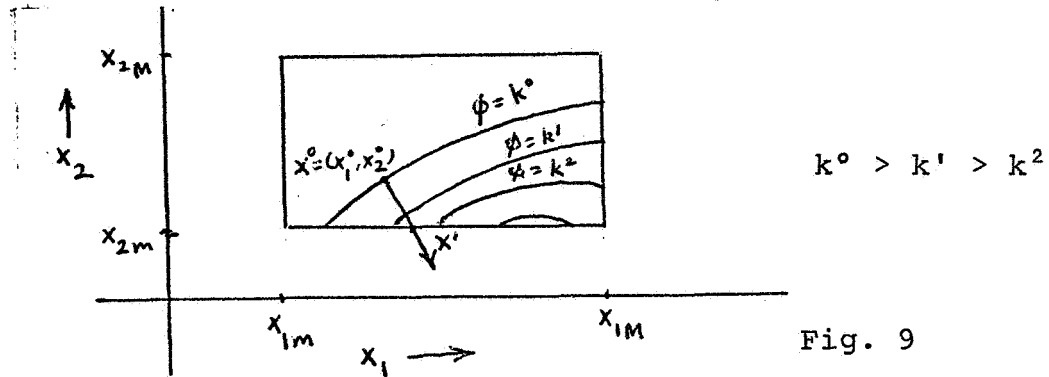


Fig. 9 shows the contours of the level lines for constant ϕ in space (x_1, x_2) . The variables x_1 and x_2 are bounded. We have to seek a minimum of ϕ . Let $x^0 = (x_1^0, x_2^0)$ be an arbitrary starting point. In the gradient technique one seeks to move in the direction of the negative gradient $-\nabla\phi$ which is normal to the level line $\phi=k^0$ at x^0 . The step size is estimated from the desired improvement $d\phi$. If x^0 is close to the boundary of x_1 or x_2 the step in the negative gradient direction may go past the boundary as shown in Fig. 9. One has to 'truncate' the step at x' which is a point on the boundary. It is apparent that from this point on, the step in the direction of the negative gradient will be truncated in the x_2 direction. The truncated step will yield much less improvement than the stipulated $d\phi$. This seriously affects the convergence.

It is obvious from the figure that the best direction to follow is

$$\begin{bmatrix} \frac{\partial \phi}{\partial x_1} \\ 0 \end{bmatrix}$$

i.e. to keep moving along the boundary $x_2 = x_{2m}$.

Block 5: Revised estimate of δu -- Let us consider the situation shown in Fig. 10 where u is a scalar function. A part of u lies on the boundary u_m and a part lies on u_M . The variations δu^ϕ and δu^Ω are the components obtained as described in Block 4. The variation δu is the first estimate. However after truncation it reduces to $\delta u'$. It is apparent that a large section of δu -- shown hatched -- was counted upon to make substantial contribution towards the variations $d\phi$ and $d\Omega$, but is now ineffective. The composition of $\delta u'$ in terms of δu^ϕ and δu^Ω cannot be estimated. Since δu^ϕ affects $d\Omega$, and δu^Ω affects $d\phi$, the corrections $d\phi$ and $d\Omega$ resulting from $\delta u'$ are not only small but are at times far different from the stipulated values.

This can be remedied, to a large extent, by giving due consideration to the effect of truncation in the revised estimate of δu . This is effected by using $|\delta u'|$ or $(\delta u')^2$ as a weighting factor. Thus we have

$$\begin{aligned} \delta u^{\phi'} &= (\delta u')^2 \delta u^\phi \\ \delta u^{\Omega'} &= (\delta u')^2 \delta u^\Omega \end{aligned} \tag{4.8}$$

Wherever the first estimate δu gets truncated, $\delta u'$ is

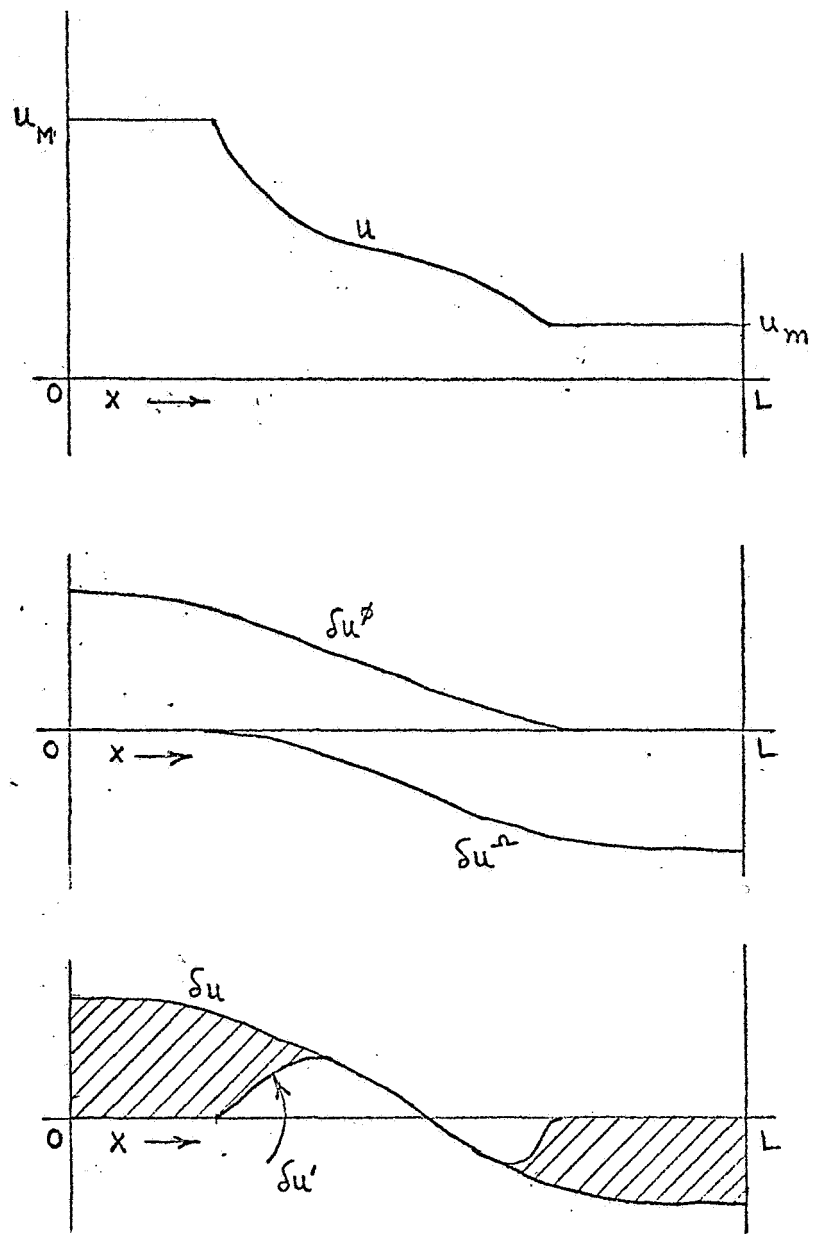


FIGURE 10

equal to zero (see Fig. 10). The new estimates δu^ϕ and δu^Ω will also be zero wherever the first estimate δu is truncated. Thus, the second estimate of the step sizes v^ϕ and v^Ω is obtained by reshaping δu^ϕ and δu^Ω so that the second estimate of the variation δu is confined, as far as possible, to the region where the possibility of the variation exists.

Revised values of v^ϕ and v^Ω may be obtained from (3.43) and (3.44) and the new estimate of δu is given by

$$\delta u = v^\phi \delta u^\phi + v^\Omega \delta u^\Omega,$$

Effectively we use a W factor, so that

$$W^{-1} = \begin{bmatrix} (\delta r') \left[\frac{(r_M - r_m)}{L} x + r_m \right] & 0 \\ 0 & (\delta c') \left[c_M - \frac{(c_M - c_m)}{L} x \right] \end{bmatrix} \quad (4.9)$$

where $\delta r'$ and $\delta c'$ are truncated first estimates from Block 4.

The last part of this operation is checking and truncating $u + \delta u$. Then the program goes back to Block 2 for the next iteration cycle.

COMPUTER SOLUTIONS

The use of the Hybrid Computer was considered to be best suited for this problem due to the following reasons:

- (i) The Analog computer can solve the differential equations without discretization in 'x' space.
- (ii) The Digital computer with the help of D to A converter can generate arbitrary shapes of distributions and feed them to the Analog computer to obtain the representation of a non-uniform transmission line.
- (iii) The storage facility and the computational capability of the Digital computer can be utilized to evaluate the estimates of δu .

Appendix A describes the Hybrid Computer operations. As a particular case of the oscillator problem we chose the following set of values for the numerical analysis.

The ratio of r_M/r_m and c_M/c_m is chosen to be 10. The limiting values are chosen to be

$$\begin{aligned} r_M &= \omega c_M = .8 \\ r_m &= \omega c_m = .08 \end{aligned}$$

This choice is governed by the limitations of the dynamic range of the system. The ADC, DAC, and analog units cannot handle quantities larger than unity (10 volts), and for the values of the order of .0010 there is a serious noise problem. However, a large spectrum of values can be handled by transforming the independent variable (thus effectively changing the scale)

provided the dynamic range is not too large.

The quantum of the interval is chosen to be 1/10 unit. Thus we have 189 discrete intervals for length of 18.9 units. The functions $r(x)$ and $\omega_c(x)$ are represented by stepwise approximation. At the start of each interval the value of $r(x)$ or $\omega_c(x)$ at that point is provided on DAC and held constant until the start of the next interval. The system variables are sampled on the ADC at the end of each interval. The sampled values are fed to the DC through ADC while the integration continues uninterrupted.

The first order unimproved gradient technique with W chosen as an identity matrix was tried first. Different distributions such as uniform distribution, ramp distribution, or exponential distribution were used as an initial guess. The problem of sensitivity was immediately felt since they did not converge to a single distribution.

The unimproved first order technique with W as an identity matrix indicated that with different initial guesses the iterations moved the distributions in the same general direction but the sensitivity problems prevented them from converging to a single distribution. Also, the simultaneous convergence of ϕ and Ω was affected when the δu estimates were truncated. (see Appendix B) The algorithm described in Chapter 4 for the improved first order technique is an attempt to correct these defects. The method seems to work satisfactorily.

The unimproved first order technique with W as an identity

matrix is used until the improvement in ϕ becomes small. Then we branch to the method using (4.9) as weighting factor, which is the improved gradient technique.

The optimal distributions obtained from the computer are quite noisy. (The reasons are described in the next chapter.) Fig. 11 is the noisy output from the computer. Fig. 12 gives the filtered version. The rest of the figures presented here are the filtered versions of the computer output.

In order to check the dependence of the final distributions on the initial guess, two widely different sets of distributions are selected as an initial guess. In each case the length of the line is 14.

Case 1: The initial distributions are

$$r(x) = \omega c(x) = 0.325$$

The final distributions are given in Fig. 12.

Case 2: The initial distributions are

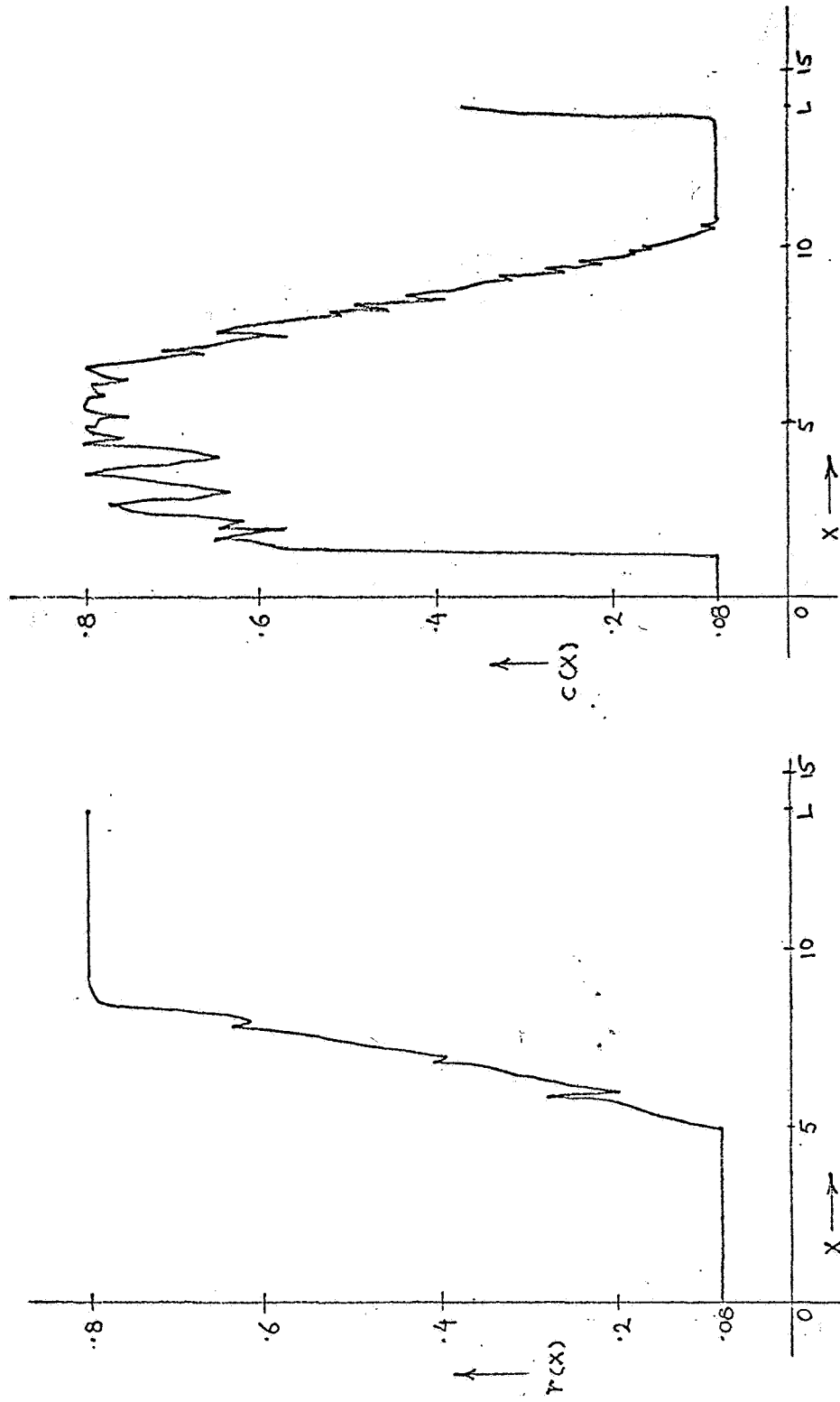
$$r(x) = 0.8 - \left(\frac{0.8-0.08}{L}\right)x$$

$$\omega c(x) = 0.08 + \left(\frac{0.8-0.08}{L}\right)x$$

The final distributions are given in Fig. 13

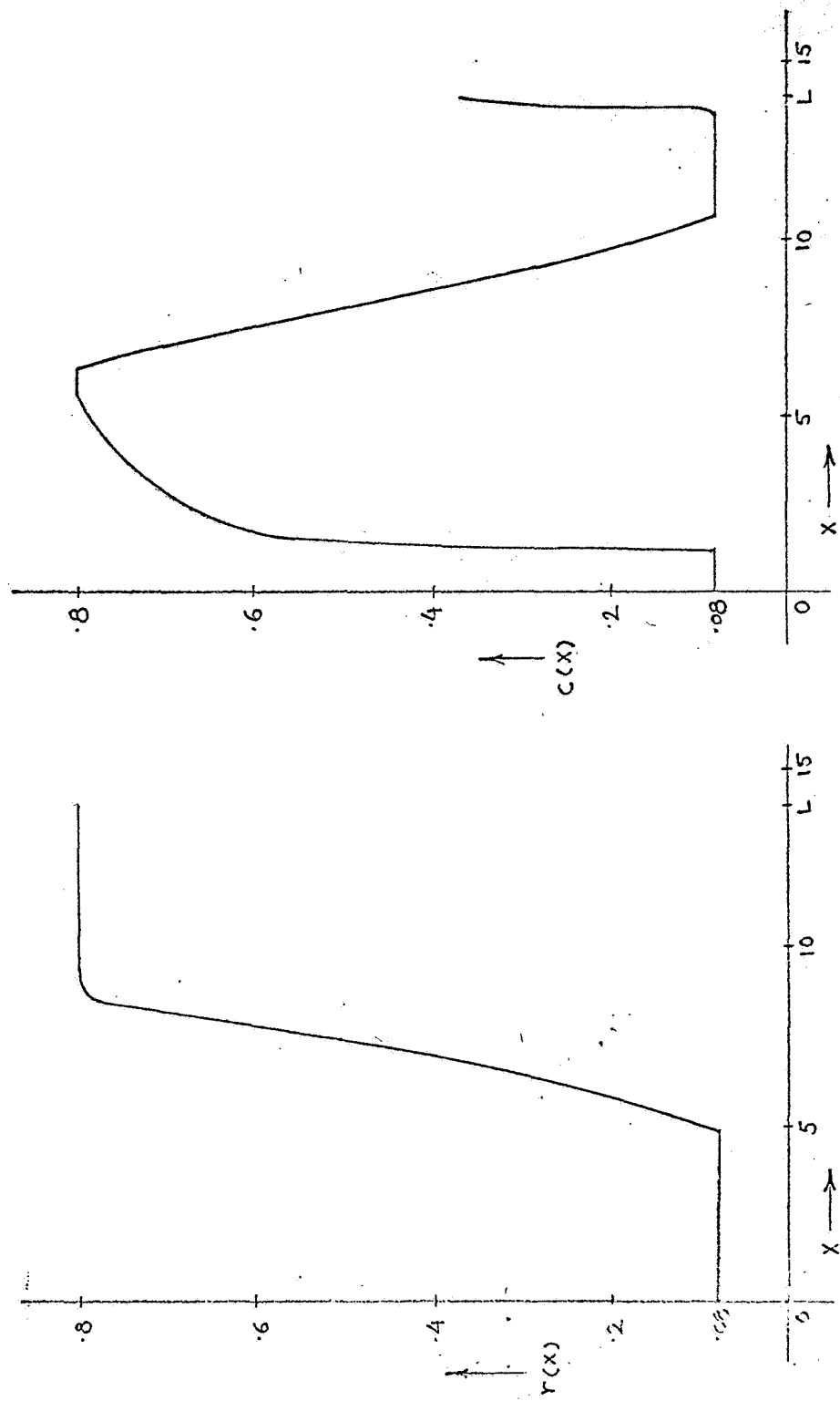
The comparison of the results recorded in Fig. 12 and Fig. 13 shows that in both cases the distributions converged to the same set of final distributions. This indicates that the algorithm derived here is quite insensitive to the choice of the initial distributions.

On the optimal switching curve in between the boundaries, H_u should be identically zero. The observed values of H_u at



Unfiltered Final Distributions, $L=14$, $\lambda(x)=0$

FIGURE 11



Filtered Final Distributions, $L=14$, $l(x)=0$

FIGURE 12

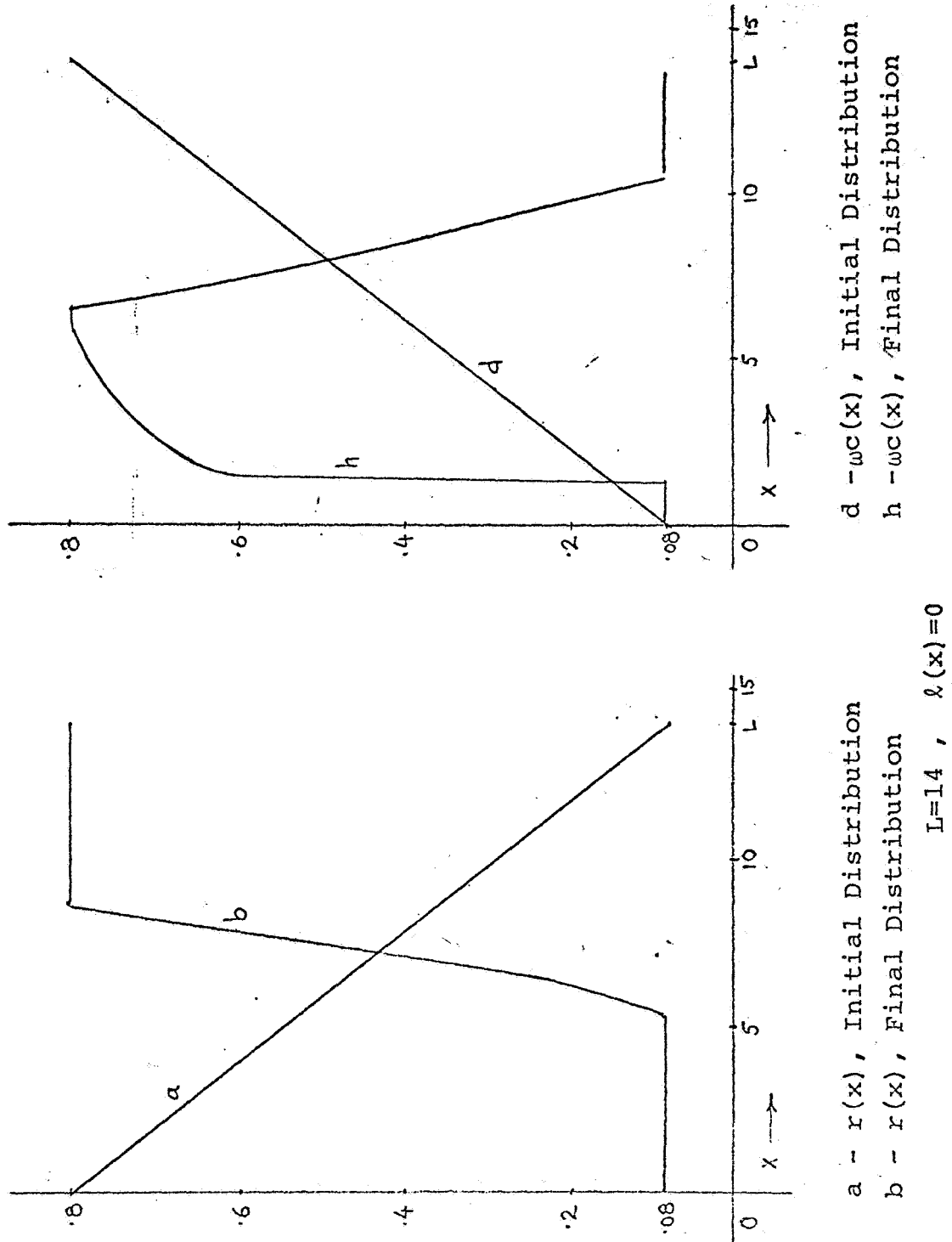


FIGURE 13

the start of the iteration and at the end differed by a factor of about 1000 indicating that we are very close to the optimum.

Fig. 14 through 18 give the results for different assumed length 'L'. In each case the starting distributions are taken to be uniform and inductance $\ell(x)=0$. Table I summarizes these results.

For the second set of results we assumed different values for inductance $\ell(x)$. As stated before $\ell(x)$ was assumed to be non-controllable and constant.

Figures 19 through 21 present the optimal $r(x)$ and $c(x)$ for different $\ell(x)$. Table II summarizes these results.

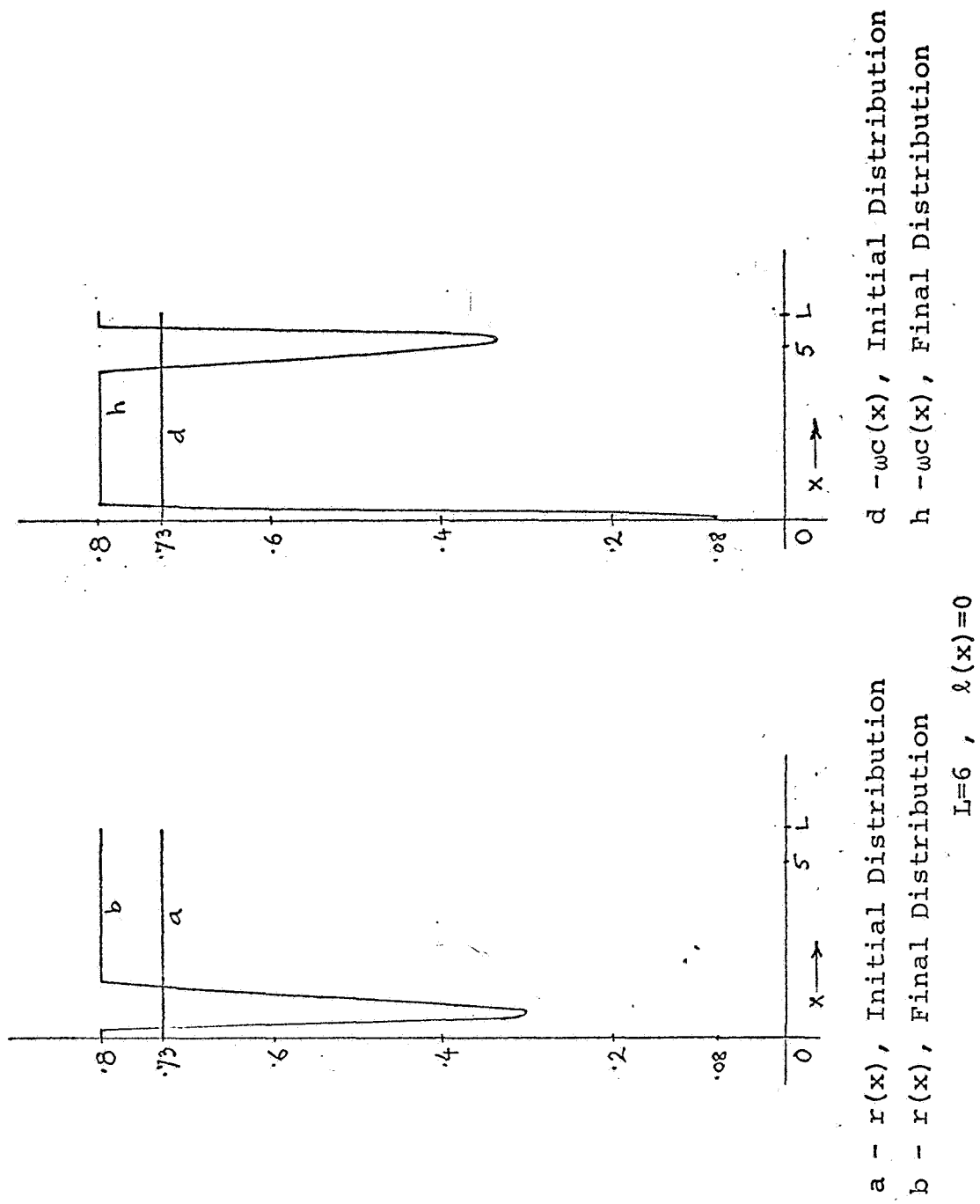


FIGURE 14

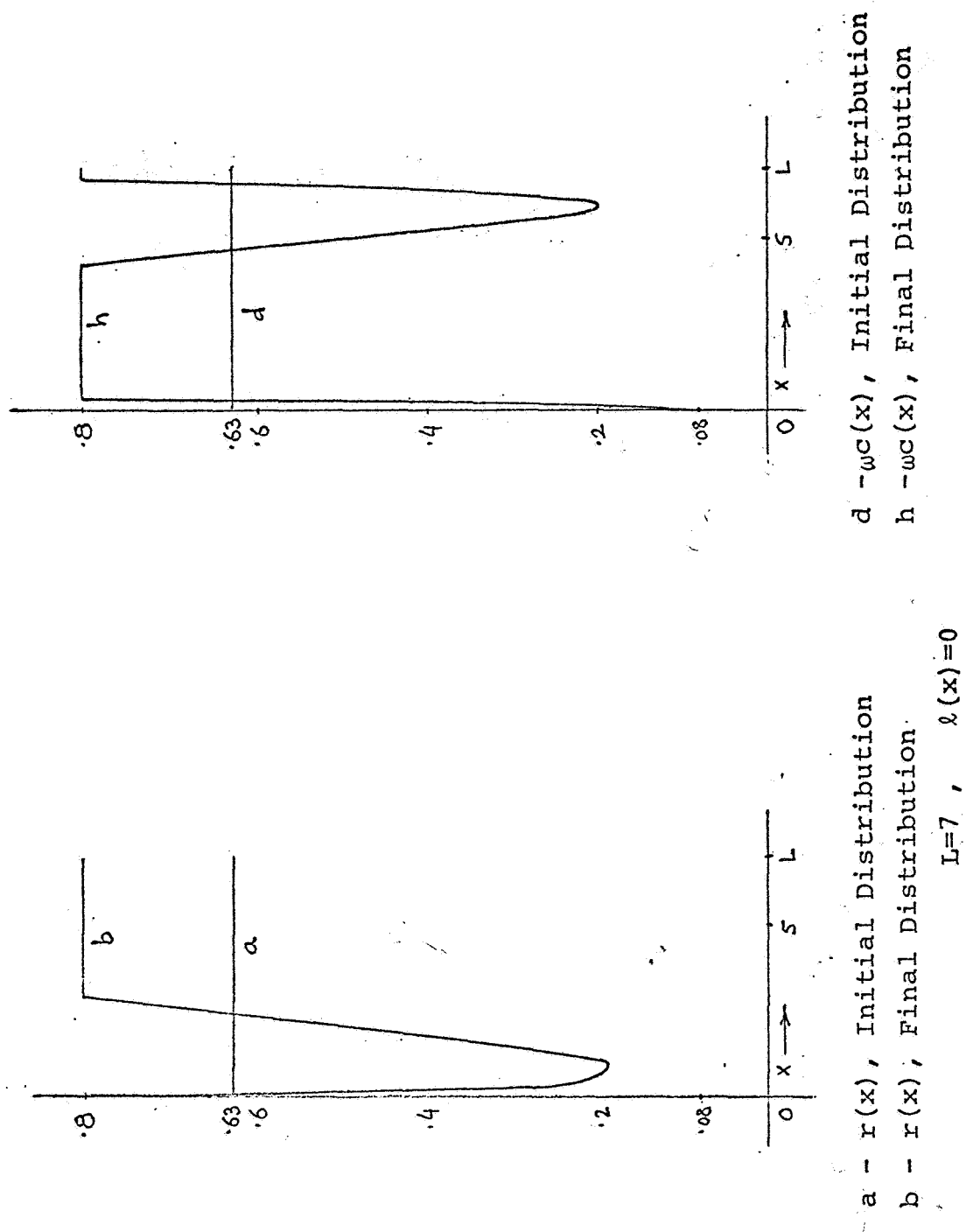
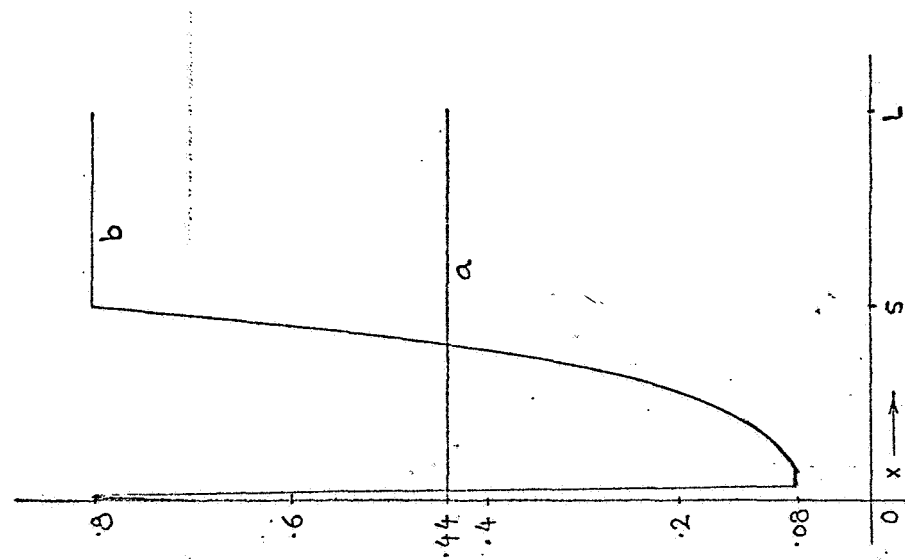
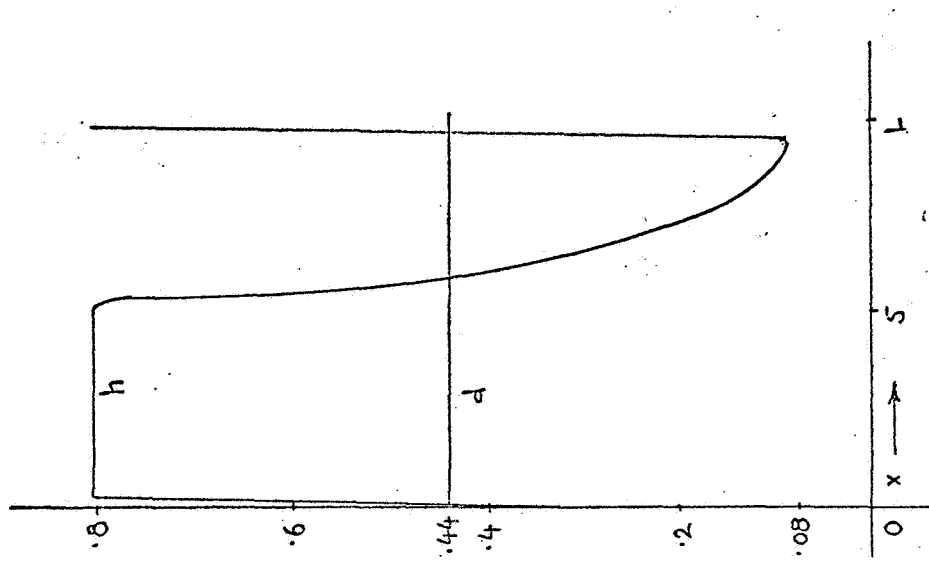


FIGURE 15



a - $r(x)$, Initial Distribution
 b - $r(x)$, Final Distribution

$$L=9.9, \quad \ell(x)=0$$



d - $wc(x)$, Initial Distribution
 h - $wc(x)$, Final Distribution

FIGURE 16

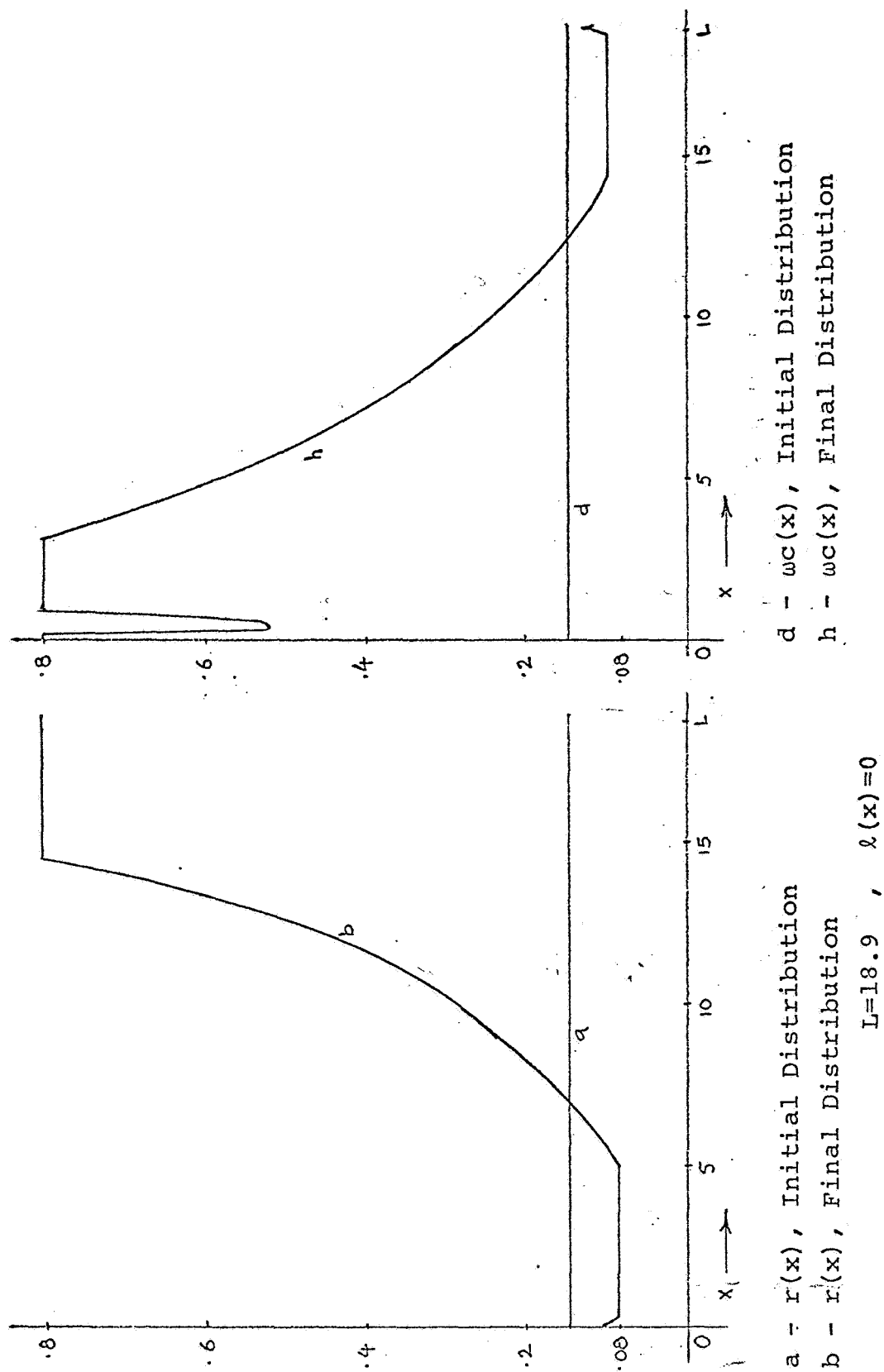
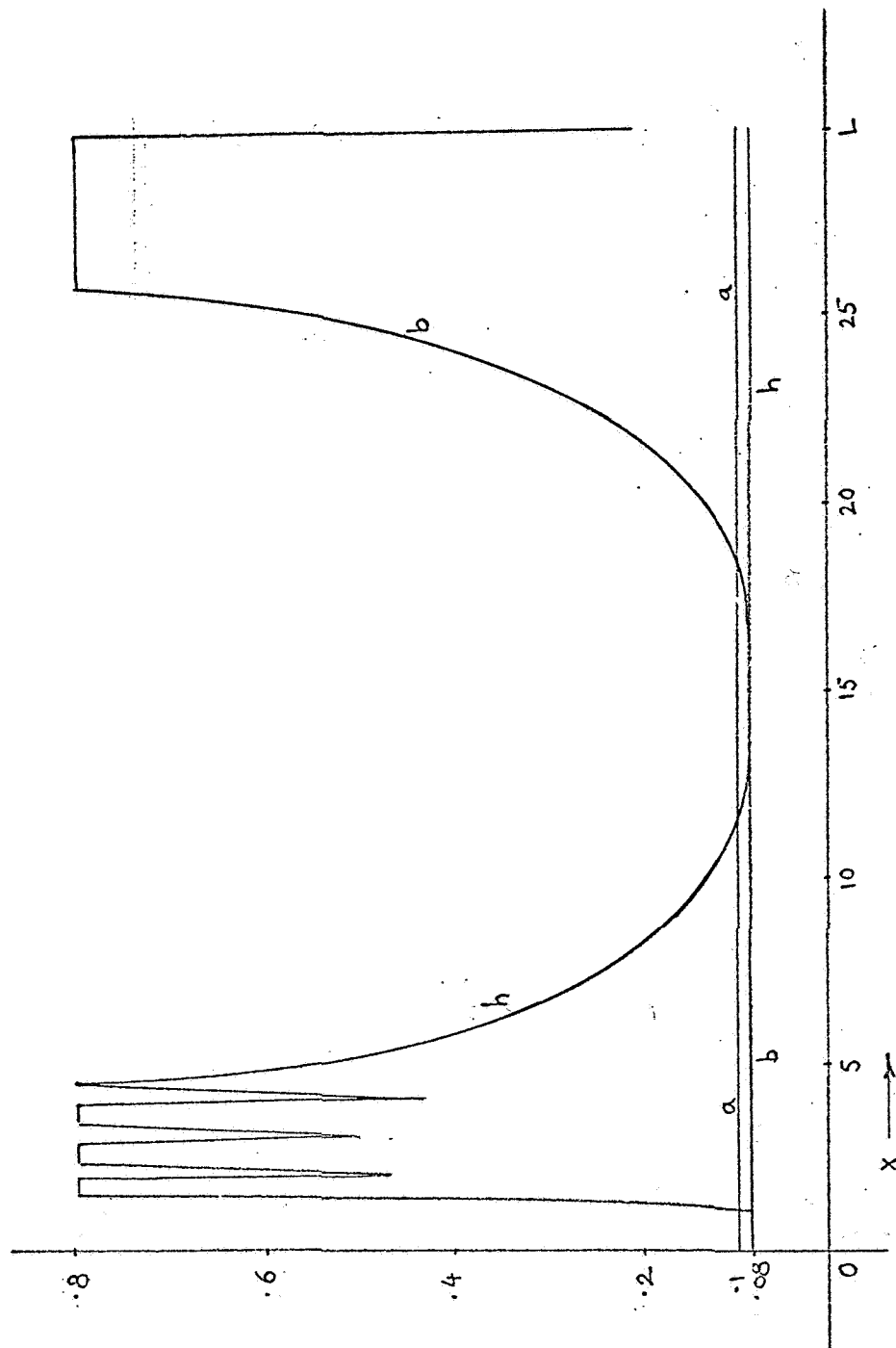


FIGURE 17



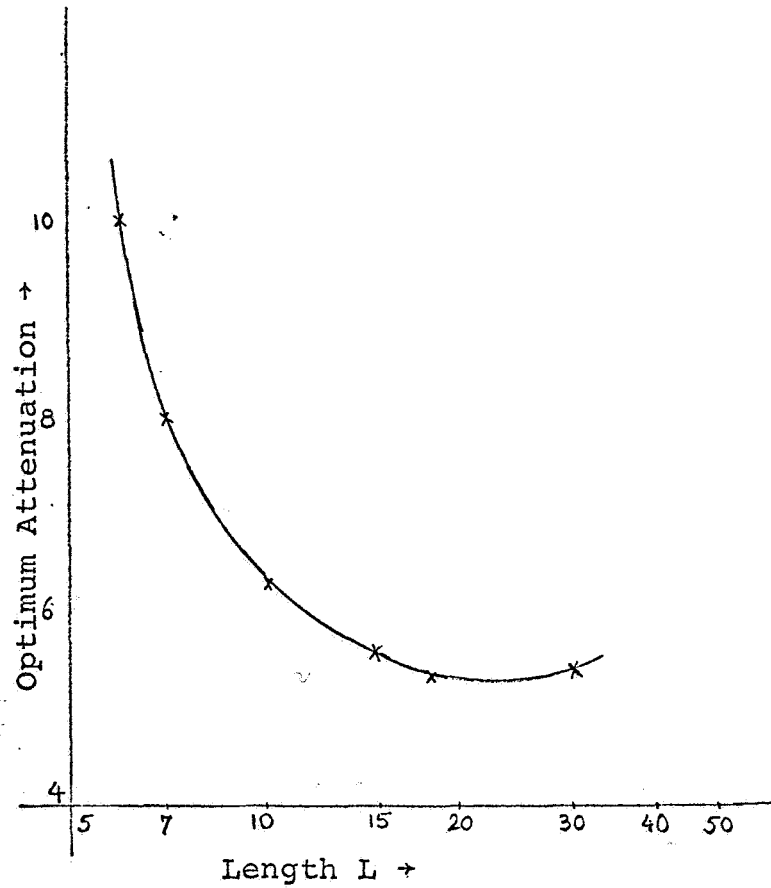
a - $r(x)$ and $\omega(x)$, Initial Distribution

b - $r(x)$, Final Distribution

h - $\omega(x)$, Final Distribution

$L=300$, $\ell(x)=0$

FIGURE 18



Optimum attenuation as a function of the total length of a line.

FIGURE 18a

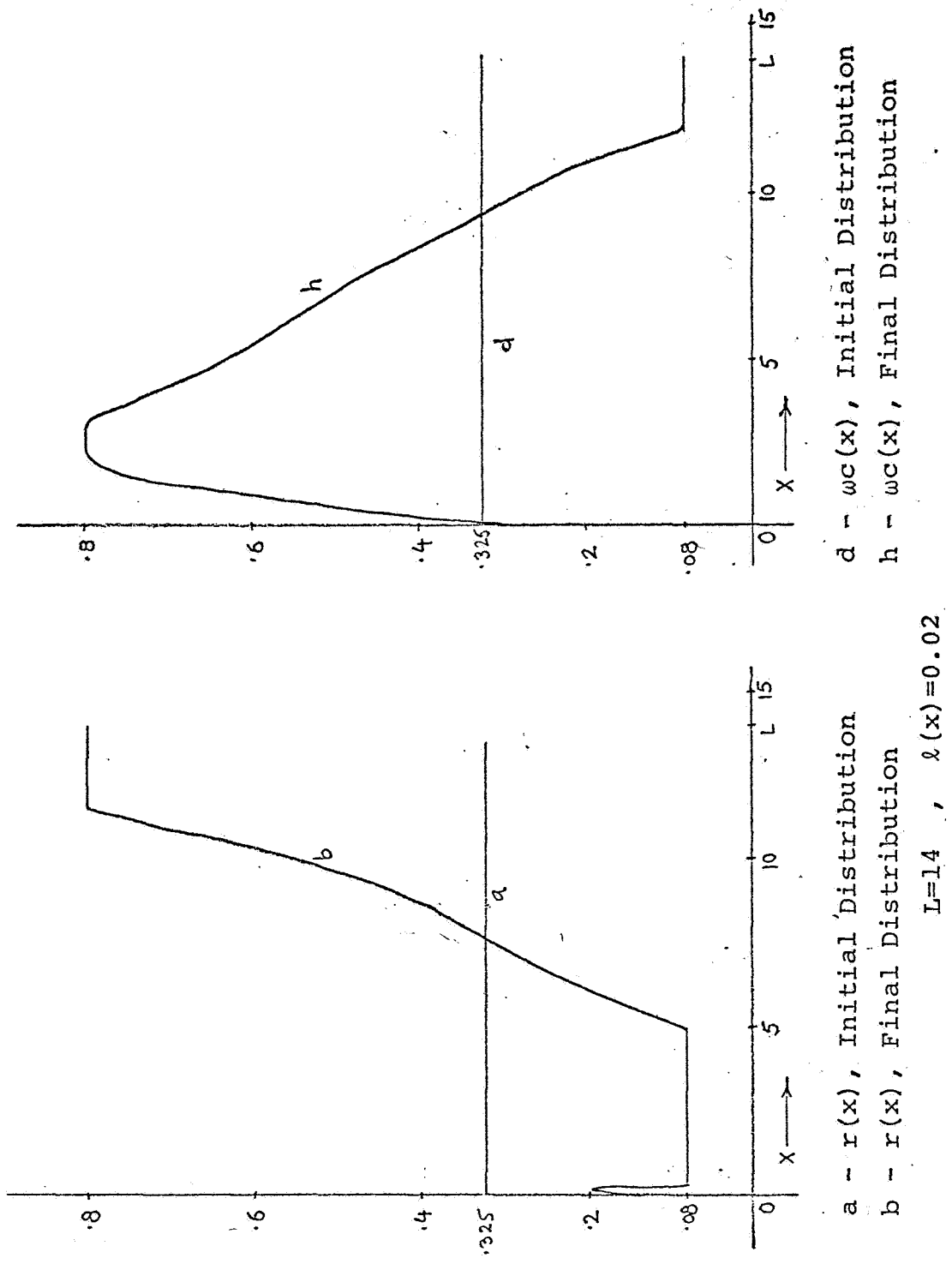


FIGURE 19

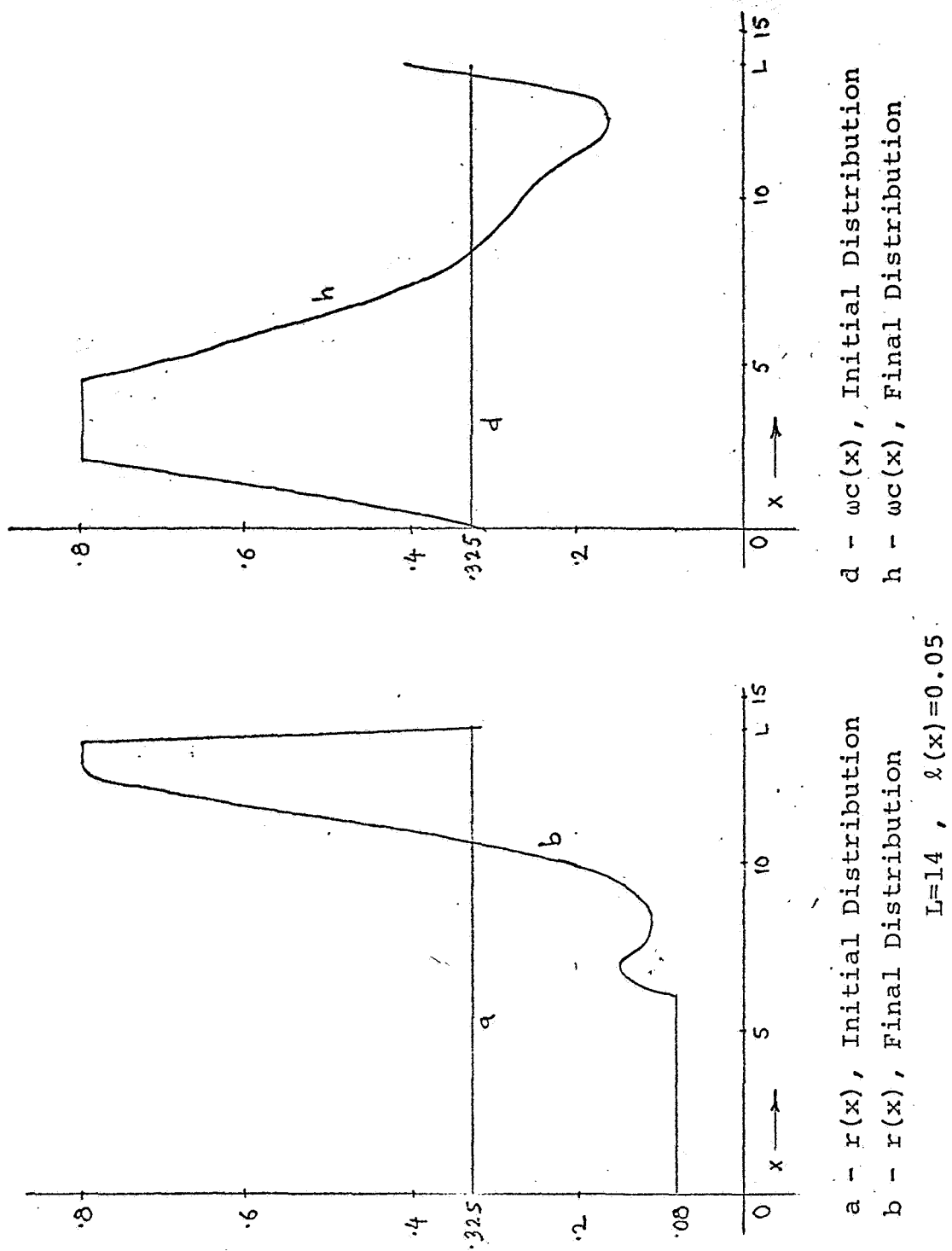
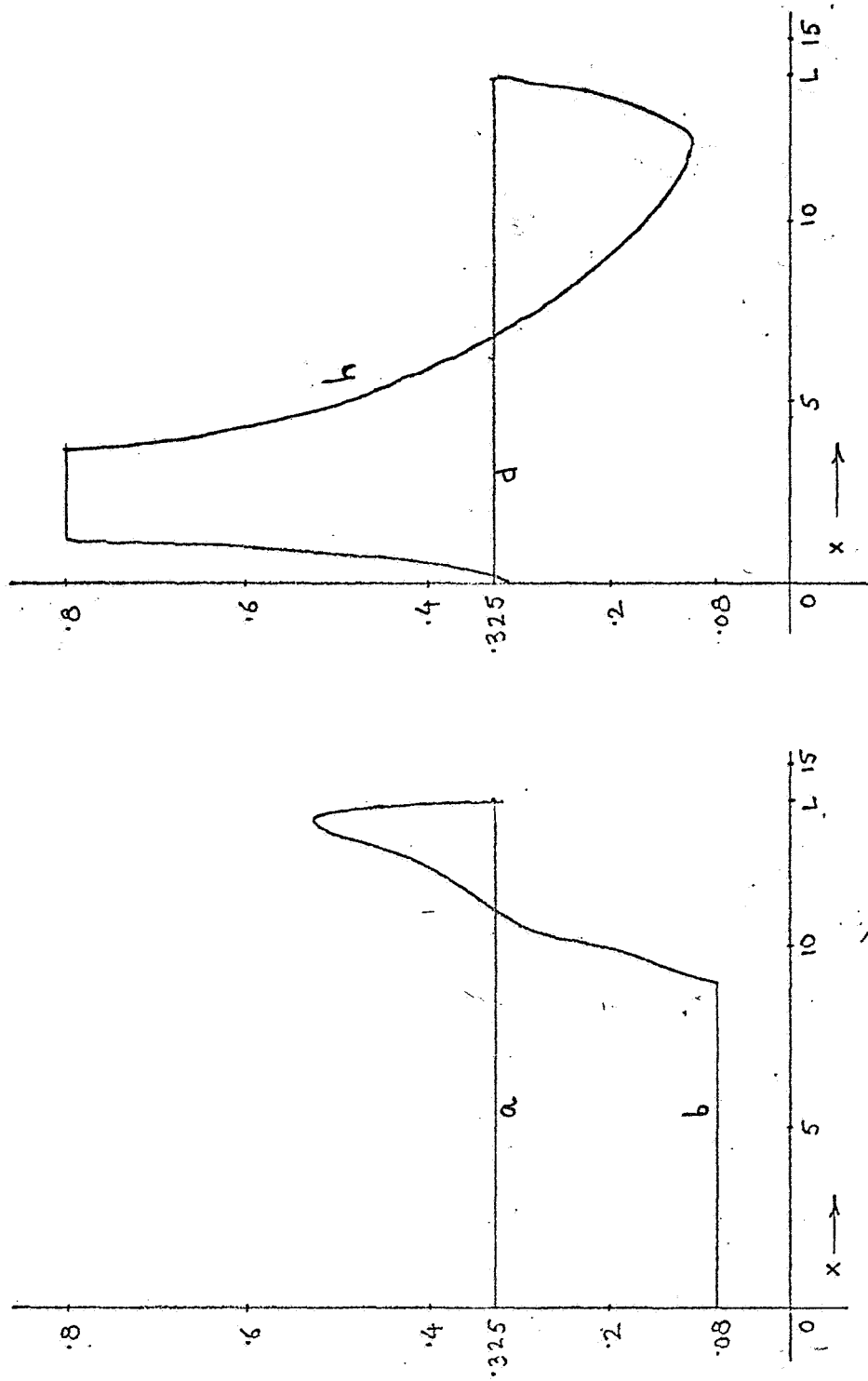


FIGURE 20



a - $r(x)$, Initial Distribution d - $\omega c(x)$, Initial Distribution
 b - $r(x)$, Final Distribution h - $\omega c(x)$, Final Distribution

$L=14$, $\lambda(x)=0.1$

FIGURE 21

ERRORS AND LIMITATIONS

A. Scale and Range

The analog computer is a 10 volt machine. The DAC is a 10 volt unit with 14 bits plus a sign bit and the ADC is a 10 volt unit with 13 bits plus a sign bit. Thus the lowest voltage level that the setup can handle is about 2 mv, as decided upon by the ADC. Any voltage level below 2 mv is interpreted as a zero by the ADC and the voltage levels above 10 volts are either rejected by the converters or cause saturation of the amplifiers. Thus the dynamic range of the setup is 5×10^3 .

B. Noise

(i) Random Noise -- The individual component of the system has a specified noise level as given below.

ADC - the noise level is ± 1 bit, equivalent to about ± 2 mv.

DAC - the noise level is negligible as compared to that of the ADC and AC.

Analog Computer - the nonlinear multipliers have the highest noise level. It is specified to be ± 3 mv. However, when the transmission line equations were integrated a number of times using the entire Hybrid setup, for the same distributions $r(x)$, $c(x)$ and $l(x)$ the end point values of the voltage $V_1(x)$ were found to be repeatable within 20 mv.

(ii) Quantization Noise -- The ADC while reading the results from the Analog Computer quantizes them. The random noise is superposed on top of this quantized signal. In the

algorithms these readings are operated upon and amplified -- especially during the last part of the iteration -- several times. Thus 2 mv quantization step and about 6 mv noise can cause a noise level in the range of a hundred mv. Figures 12 through 21 are the smoothed out versions of the computer output. Fig. 11 is one of the original computer output.

The noise problem becomes more serious with the complicated algorithms involving large numbers of algebraic operations. For this reason, the algorithm should be as simple as possible.

C. Limitations of the Method

H_u is a smoothly varying function. Thus every variation in the control has a continuous first derivative in the open region. If the optimal distribution has a discontinuous first derivative and the initial estimate does not, the solution will not converge on to the optimal. Also if the initial guess has a discontinuous first derivative we can never get rid of this discontinuity in the open region. In the present case the uniform, ramp, exponential distributions all converged to the same distribution. However, when the initial guess was a bang bang type of distribution, the final distribution retained the kinks.

CONCLUSIONS

1. With the bounds on resistance and capacitance decided upon by the fabrication limitations, and the length prespecified, the optimum 180° phase shift network with minimum attenuation turns out to have distributions of r and c that have limiting values with the singular switching curves.

The attenuation of unity, as projected by Johnson and calculated from Edson's results is not realizable due to the physical limitations.

The optimum attenuation is not far better than what can be achieved by exponential distributions given a free choice of length.

2. It is possible to obtain a solution to a 'singular' optimization problem by using the Improved Gradient Technique developed here.

APPENDIX A

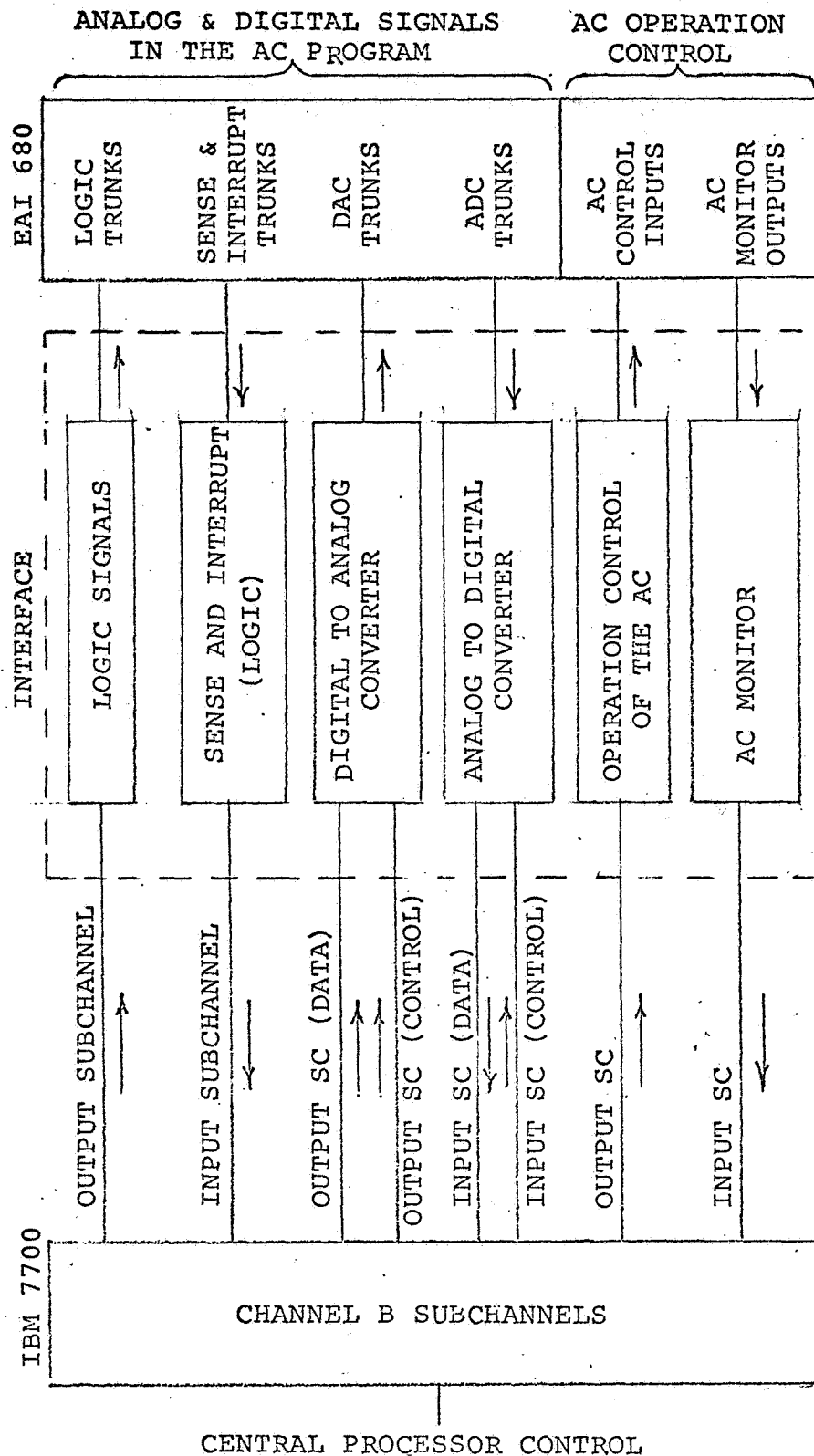
Hybrid Computer

This is a combination of the Analog and Digital computers. We have EAI680 analog computer and IBM7700 digital computer with input-output subchannels for the transfer of the information. In order to transform this setup into a hybrid unit, we designed and built the interface. Fig. A.1 shows the flow diagram for the hybrid unit.

(i) Digital computer: The DC contains the multiplexor channel, channel B. It permits the attachment of different data acquisition and data distribution devices to the processor of the DC. The input subchannels of channel B are capable of recording the logic levels -- true or false -- of the incoming lines and the output subchannels can send the desired logic levels on the output lines. The operation of channel B is controlled by the central processor unit.

(ii) Interface: The interface provides the medium of communication between the AC and DC. It is essentially a translator unit. The function of the various sections of the interface are described below.

Operation Control of the AC: The operation of the AC is controlled by the coded logic signals sent from the DC. The interface converts the input logic levels into the appropriate output logic levels and also generates the clock pulses required for certain operations.



The arrows indicate the direction of the flow of the information.

FIGURE A-1

The operations controlled are as follows:

- 1) Operate (Integrate), Hold, Initial condition, etc.
- 2) Analog component selection for readout or potset;
e.g., Amplifier, Trunk, Pot, etc.
- 3) Time constant selection, e.g. Seconds, Milliseconds, etc.
- 4) Digital mode selection, e.g. Set, Clear (Registers,
Counter), etc.
- 5) Digital clock rate selection
- 6) Selecting the address of the analog component
- 7) Setting a pot coefficient

AC Monitor: The coded logic signals coming from the monitor of the AC are transmitted to DC. The DC compares the control order with the monitor signal to find out whether the execution is proper.

Logic Signals: Certain decisions made by the DC regarding the status of the program under execution are transmitted through interface to the logic trunks. These signals can be used to effect a change in the AC program.

Sense and Interrupt: The status of the AC program such as a comparator output is conveyed to the interface on the sense lines. The interface in turn transmits the message to the DC. The interrupt lines are used for conveying the undesirable status of operation such as overload. The AC is programmed to interrupt the operation under such conditions.

Digital to Analog Converter: This is an eight channel serial input, parallel output unit. The control signal from

the AC initiates the conversion of the digital data on the input lines from the DC into the analog signal. The analog signal appears on the channel selected by the control word from the DC. The output channels are connected to the DAC trunks on the AC.

Analog to Digital Converter: This is a 24 channel parallel input serial output unit. It receives the analog input from the ADC trunks. The control word from the DC selects the channel and initiates the conversion. The digital output is transmitted to the DC.

(iii) Analog Computer: The AC can be divided into three sections.

Analog Section: It consists of the analog components such as integrators, summing amplifiers, track & store amplifiers, etc. ADC trunks receive the inputs from this section and DAC trunks supply the analog signals to this section.

Logic Section: This section contains the logic elements such as gates, counters, registers along with the clock outputs and control inputs for certain analog components. The sense and interrupt trunks receive the inputs from this section. The logic trunks appear in this section.

Operation Control: This section controls the operation of both the analog and logic sections. It controls all of the operations listed under "Operation Control of the AC" in the description of the interface. It receives the coded control word, either from pushbuttons or from the interface. It also

generates the monitor signals.

Hybrid Operations

The two important links in the hybrid setup are the DAC and the ADC.

DAC: The output subchannel of the DC transmits the digitized value of the variable. The load command from the DC loads the word into the registers of the DAC. However, unless the DAC channel receives the enable command the analog output does not appear at the output terminal of the DAC. The previous value is retained at the output until a new enable command is received.

ADC: The DC selects the ADC channel by controlling the multiplexor switches. The conversion of the analog signal on this preselected channel is initiated by the start pulse. On completion of the conversion a pulse is sent to the input subchannel of the DC. On receiving this pulse the input subchannel registers the digital output of the ADC. This is subsequently transferred to the memory of the DC.

Setting up initial conditions and static test:

The operation control subroutine sets the AC in the "set pot" mode. The proper address word selects the desired servo controlled pot. The value register is loaded and the servo start pulse transmitted from the DC. The monitor subroutine checks if the proper pot has been selected and the operation completed. Thus the initial condition -IC- is established with the help of servoset pots.

The AC is then driven into the IC mode and outputs of amplifiers are read on the ADC. This gives the static test.

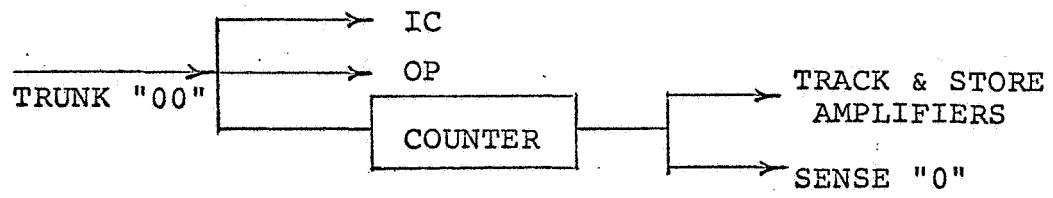
Integration routine: A subchannel of the DC is used for starting and terminating the integration operation. Selection of the counter SC turns trunk "00" (Fig. A.2) on and the AC goes into "operate" mode thus starting integration. At the same time, the AC counter starts counting AC clock pulses and gives the output as in Fig. A.2. The monostable multivibrator (Fig. A.3) generates a pulse every 1000 μ sec. which generates a DC interrupt. The DC counts the number of such interrupts. As soon as the DC counts a specified number of pulses it deselects the subchannel terminating the integration operation and driving the AC into the IC mode. The pulse from the monostable multivibrator also starts the conversion and enables the DAC channels.

Before the start of integration:

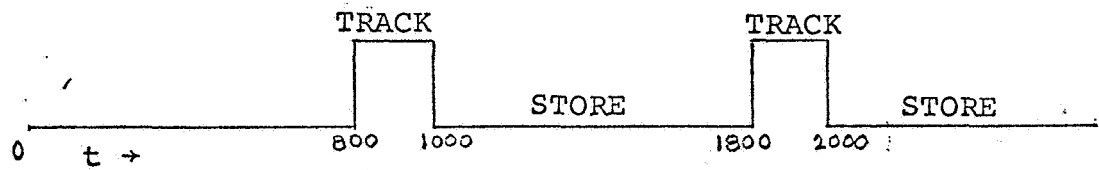
- (i) AC counter is reset,
- (ii) AC clock mode is selected (such as 10 kc, 100 kc, 1000 kc),
- (iii) AC time constant is selected (such as seconds, milliseconds, etc.),
- (iv) The values of DAC functions for the second interval are loaded.

(For the oscillator problem the clock mode was 1000 kc and the time constant was 0.1 sec.)

Now the integration is started by selecting the counter SC. Fig. A.4 describes the flow of events.



LOGIC PATCHING OF AC



AC COUNTER OUTPUT

FIGURE A-2

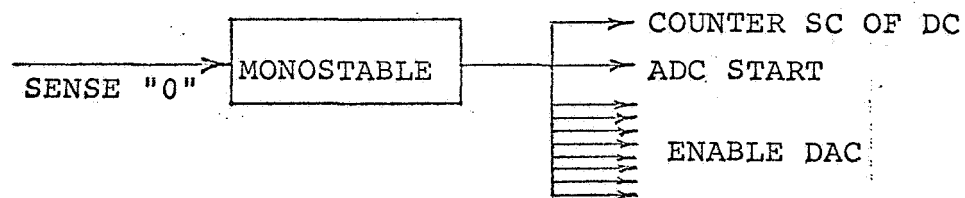
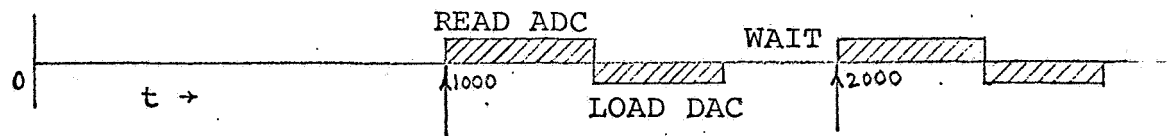


FIGURE A-3



SENSE "0" PULSE

FIGURE A-4

As the first counter pulse comes in, it enables all the DAC channels. Thus, values of all the co-efficients for the second interval are made available. All the track and store amplifiers go into store mode thus preserving the values at the instant of the counter pulse. The DC now selects and reads the ADC channels one by one. This is followed by serial loading of DAC channels with the values for the next interval. This completes the operations for one interval and the DC waits for next counter pulse. The process repeats until the counter SC is deselected.

APPENDIX B

For a system represented by

$$\frac{d}{dx} y = f(y, u, x) ,$$

with criterion function $\phi(y(0), y(L))$, the functional relationship between a variation in ϕ and the variation ' δu ' in control u is obtained as (ref. equation (3.37))

$$d\phi = \int_0^L H_u \delta u dx . \quad (B.1)$$

Let us assume that ϕ is to be maximized. In the Gradient Technique the hope that the iterations would converge is based on obtaining a positive $d\phi$ as a result of every iteration cycle. Thus we can stipulate three necessary conditions for $\delta u(x)$,

(i) $\text{Sgn } \delta u(x) = \text{Sgn } H_u(x)$, for a finite length and $\delta u(x) = 0$ for the rest of x . This assures $d\phi \geq 0$, (B.2)

(ii) $u_m \leq u(x) + \delta u(x) \leq u_M$,

(iii) $\int_0^L \delta u(x), \delta u(x) dx \ll 1$.

This assures that the variation $u(x)$ is small enough to justify the first order approximations made in the derivation of (3.37)

Let us define $u_m - u(x) = \delta u_m(x)$ and $u_M - u(x) = \delta u_M(x)$.

The bounds on $\delta u(x)$ can now be specified as

$$\delta u_m \leq \delta u(x) \leq \delta u_M ,$$

Since $u(x)$ is an admissible control vector

$$u_m \leq u(x) \leq u_M .$$

Hence

$$\delta u_m(x) < 0 ,$$

and

$$\delta u_M(x) > 0 .$$

Let $\delta u(x)$ be any function that satisfies the first and the last condition stated in (B.2). (See Fig. B.1) The function δu can be expressed as a sum of a function δu_p and δu_n such that

$$\delta u_p(x) \geq 0 ,$$

$$\delta u_n(x) \leq 0 ,$$

and $\delta u = \delta u_p + \delta u_n .$

The variations δu_p and δu_n also satisfy the first and the last conditions stated in (B.2). The functions $\delta u_p(x)$ and $\delta u_n(x)$ can be further divided so that

$$\delta u_p(x) = \delta u_{pa}(x) + \delta u_{pt}(x) ,$$

$$\delta u_n(x) = \delta u_{na}(x) + \delta u_{nt}(x) ,$$

where $\delta u_{pt}(x)$ and $\delta u_{nt}(x)$ are the truncated sections of δu_p and δu_n respectively.

We have $\text{Sgn } \delta u_p(x) = \text{Sgn } \delta u_{pa}(x)$

and $\text{Sgn } \delta u_n(x) = \text{Sgn } \delta u_{na}(x)$

Thus $\delta u_{pa}(x)$ and $\delta u_{na}(x)$ satisfies the first and the last condition stated in (B.2). They also satisfy the second condition (See Fig. B.1).

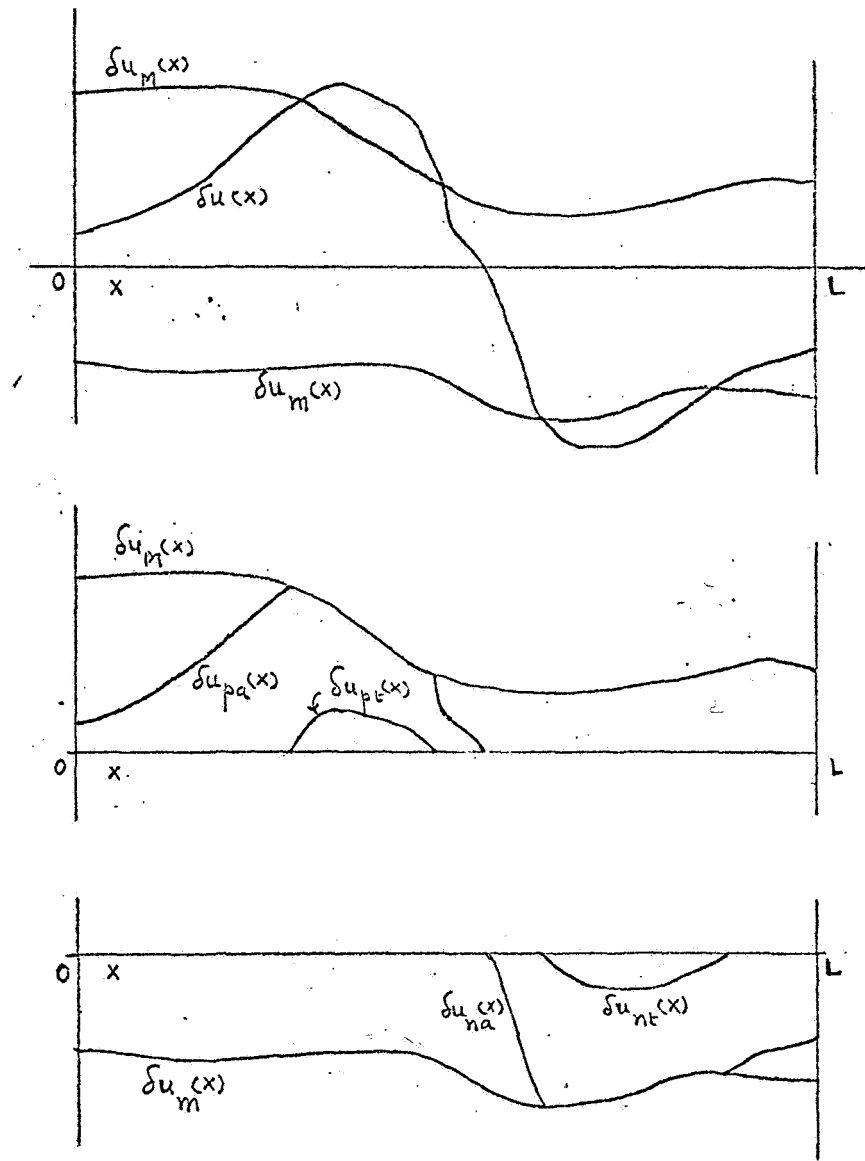


FIGURE B-1

The same is true about δu_a where

$$\delta u_a(x) = \delta u_{pa}(x) + \delta u_{na}(x) \quad (B.3)$$

The function $\delta u_a(x)$ is a truncated part of $\delta u(x)$. Hence the truncation does not violate the conditions for convergence of the Gradient Method.

However with more than one target function, such as ϕ and Ω , δu is composed of more than one component such as

$$\delta u = v^\phi \delta u^\phi + v^\Omega \delta u^\Omega \quad (B.4)$$

and the functional relationship is (See (3.43)).

$$d\phi = v^\phi \int_0^L H_u^\phi \delta u^\phi dx + v^\Omega \int_0^L H_u^\Omega \delta u^\Omega dx \quad (B.5)$$

In such a case δu^Ω affects $d\phi$ (and δu^ϕ affect $d\Omega$). The condition (i) holds true for the first term on the R.H.S. of equation (B.5). However, the second term does not necessarily satisfy the condition (i). Besides, δu^ϕ and δu^Ω are not truncated separately. The truncation of δu does not provide any information as to how the truncation affects the components δu^ϕ and δu^Ω . Thus the argument about convergence breaks down.

It is observed during the numerical calculations on computer that before the control distributions reach the limiting values the first order gradient technique (using first estimate of δu) yields improvement in both ϕ and Ω simultaneously. However, once the control variables reach the boundary only one of the two improves and the other starts deteriorating. Thus a simultaneous convergence breaks down.

TABLE 1

Total length of line 'L'	Uniform distribution for 180 degree phase shift $\omega c(x) = r(x)$	Optimum attenuation
60	.73	10
70	.63	8
99	.44	6.3
140	.32	5.6
189	.235	5.3
300	.15	5.4

TABLE 2

Total length of a line 'L' = 140

Uniform line inductance $\omega l(x)$	Optimum attenuation
0	5.6
.02	4.2
.05	2.15
.10	1.15

In each case $0.8 \leq r(x)$, $\omega c(x) \leq 0.08$.

REFERENCES

1. E.L. Ginzton and L.M. Hollingsworth, "Phase-Shift Oscillator", Proc. I.R.E., February 1941, pp. 43-49.
2. R.W. Johnson, "Extending the Frequency Range of the Phase Shift Oscillator", Proc. I.R.E., September 1945, pp. 597-603.
3. W.A. Edson, "Tapered Distributed R-C Lines for Phase-Shift Oscillators", Proc. I.R.E., June 1961, pp. 1021-1024.
4. C.D. Johnson and J.E. Gibson, "Singular Solutions in Problems of Optimal Control", IEEE Trans., January 1963, pp. 4-15.
5. A.E. Bryson, W.F. Denham, F.J. Carroll, and M.K. Mikami, "Determination of Lift or Drag Programs to Minimize Reentry Heating", J. of Aerospace Sciences, 1962, pp. 420-430.
6. T.E. Bullock, "Optimization and Control Using Second Variation", Fourth Winter Institute on Advanced Control, February 20-26, 1967, Gainesville, Florida.
7. J. Burchfiel and M. Athans, "Design of Waveguides and Transmission Lines by the Distributed Maximum Principle", Proceedings of the Second Annual Princeton Conference on Information Sciences and Systems, 1968, pp. 257-263.
8. S.M. Roberts, "Dynamic Programming in Chemical Engineering and Process Control", Academic Press, 1964, pp. 110-114.